



Yu, G., He, L., Rea-Dickins, P., Kiely, R., lu, Y., Zhang, J., ... Fang, L. (2018). Preparing for the Speaking Tasks of the *TOEFL iBT®* Test: An Investigation of the Journeys of Chinese Test Takers. *ETS Research Report Series*, 2017(1), 1-59. <https://doi.org/10.1002/ets2.12145>

Peer reviewed version

Link to published version (if available):  
[10.1002/ets2.12145](https://doi.org/10.1002/ets2.12145)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://onlinelibrary.wiley.com/doi/full/10.1002/ets2.12145> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms>

**Preparing for TOEFL® iBT Speaking tasks**

*– An investigation of the iBT-ed journeys of Chinese test-takers*

Guoxing Yu<sup>a</sup>, Lianzhen He<sup>b</sup>, Pauline Rea-Dickins<sup>a</sup>, Richard Kiely<sup>c</sup>,

Yanbin Lu<sup>d</sup>, Jing Zhang<sup>e</sup>, Yan Zhang<sup>f</sup>, Shasha Xu<sup>g</sup>, Lin Fang<sup>h</sup>

a. University of Bristol, UK

b. Zhejiang University, China

c. University of Southampton, UK

d. Tsinghua University, China

e. East China Normal University, China

f. Nanjing Tech University, China

g. Zhejiang University of Finance and Economics, China

h. Shantou University, China

## **Foreword**

(to add by ETS)

## Abstract

Language test preparation has often been studied within the consequential validity framework in relation to ethics, equity, fairness and washback of assessment. The use of independent and integrated speaking tasks in TOEFL iBT represents a significant development and innovation in assessing speaking ability in academic contexts. Integrated tasks that involve synthesizing and summarizing information presented in reading and listening materials have the potential to generate new test preparation strategies. This study investigated the experiences of over 1500 Chinese test takers and 23 teachers who were preparing for the TOEFL iBT speaking tasks. It examined the frequency of use of a number of different test preparation activities and materials, reasons and expectations for taking preparation courses, and the features of preparation courses. In addition, we examined the usefulness of test preparation from two perspectives: students and teachers perceptions, and the relationship between test preparation and performance. Data were collected via questionnaires, focus-group discussions, interviews with test takers and teachers, and classroom observations. The data showed that (i) test preparation was a hugely complex multiple-components construct, and teaching and learning test-taking strategies was the most prominent feature of intensive preparation courses, (ii) there were **significant age-related** differences in students' preparation activities and focuses, **though with small effect sizes**, (iii) there **was a** high agreement between teachers and students in their views on the usefulness of test preparation activities, **and** (iv) there existed only a weak relationship between test preparation and performance. The only significant predictor of students' test performance was the frequency of their use of the TOEFL Practice Online (TPO). The findings of the study can enhance our understanding of the pedagogic practices which characterise test preparation programmes, and contribute to the ongoing validity argument for the TOEFL iBT speaking test. **The implications of the findings for test publishers, test takers, teachers and test preparation schools are discussed with reference to the instructional, learning and affective aspects of the multi-faceted construct of test preparation.**

## Acknowledgements

This study was funded by TOEFL Committee of Examiners Research Program at Educational Testing Service. Any opinions, findings, conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of Educational Testing Service.

We gratefully acknowledge the assistance of the people who contributed to this study in different roles at various stages. Without the enthusiastic support of the administrative teams, students and teachers of the participating test preparation schools, we simply would not have **completed this study**.

## Table of Contents

Foreword.....	i
Abstract.....	ii
Acknowledgements.....	iii
List of Tables.....	vi
List of Figures.....	vii
Literature Review .....	3
TOEFL iBT Speaking .....	3
Test preparation.....	4
Defining test preparation.....	4
Preparing for IELTS and TOEFL: Process and effect.....	5
Role of test preparation materials .....	9
Research gaps and the way forward.....	10
Research Questions.....	11
Method.....	12
Participants: test takers and teachers.....	12
Instruments and procedure.....	12
Understanding the context of TOEFL iBT test preparation .....	12
Test taker and teacher questionnaires.....	13
Interviews and classroom observations.....	16
TOEFL iBT test scores.....	17
Data analysis .....	17
Results.....	19
Characteristics of student participants: Age, gender and educational qualifications .....	19
Preparing for TOEFL iBT Speaking tasks .....	20
Time management and commitment.....	20
Perception of the difficulty level of the Speaking tasks.....	23
Frequency and usefulness of different test preparation activities .....	24
Course-books, websites and other materials .....	29
Reasons and expectations for taking test preparation courses .....	33
Features of test preparation courses .....	36
Test-taking strategies .....	41
Five-component test preparation: A holistic approach.....	43
Relationship between test preparation and Speaking test performance .....	44
Discussion and conclusion.....	46
RQ 1 and RQ3: How do Chinese students and teachers prepare for TOEFL iBT speaking tasks, and how useful do they think test preparation is?.....	47
RQ4: What are the relationships between test preparation and performance?.....	52

Limitation of the study and recommendations for further research .....	54
References.....	55

## List of Tables

Table 1 Rank order of the amount of time spent on the four language skills.....	21
Table 2 Percent of time spent on Speaking test.....	22
Table 3 Difficulty of the Speaking tasks .....	23
Table 4 Frequency and usefulness of seventeen test preparation activities .....	24
Table 5 Significant difference in frequency of use of test preparation activities (16-18 vs. 19-24 year olds)....	25
Table 6 Significant difference in frequency of use of test preparation activities (applying for undergraduate vs. Master programmes) .....	26
Table 7 Significant difference in usefulness of test preparation activities (female vs. male) .....	27
Table 8 Significant difference in usefulness of test preparation activities (16-18 vs. 19-24 year olds).....	28
Table 9 Significant difference in usefulness of test preparation activities (applying for undergraduate vs. Master programmes) .....	28
Table 10 Test preparation course-books (student data).....	29
Table 11 Significant difference in frequency of use of course-books (16-18 vs. 19-24 year olds).....	30
Table 12 Significant difference in frequency of use of course-books (applying for undergraduate vs. Master programmes) .....	30
Table 13 Significant difference in usefulness of course books (female vs. male) .....	30
Table 14 Significant difference in usefulness of course books (16-18 vs. 19-24 year olds).....	31
Table 15 Significant difference in usefulness of course books (applying for undergraduate vs. Master programmes) .....	31
Table 16 Test preparation course-books (teacher data).....	32
Table 17 Rotated component matrix: reasons and expectations for taking test preparation courses .....	33
Table 18 Reasons and expectations for taking test preparation courses (student data).....	34
Table 19 Reasons and expectations for taking test preparation courses (teacher data).....	35
Table 20 Number of students in a Speaking lesson.....	36
Table 21 Purpose of taking TOEFL iBT test (teacher data) .....	37
Table 22 Teachers' self-assessment of their knowledge about TOEFL iBT Speaking .....	37
Table 23 Comparison of demand of teaching TOEFL iBT with other speaking courses .....	38
Table 24 Classroom activities organised by teachers (student data) .....	39
Table 25 Classroom activities organised by teachers (teacher data, part 1).....	39
Table 26 Classroom activities organised by teachers (teacher data, part 2).....	41
Table 27 Five components of test preparation .....	43
Table 28 Component transformation matrix .....	44
Table 29 TOEFL iBT test performance.....	45
Table 30 Test preparation activities that have significant correlations with Speaking performance.....	45
Table 31 Simple linear regression.....	46



## List of Figures

Figure 1 Frequency and usefulness of test preparation activities.....	27
---	----

## Preparing for TOEFL® iBT speaking tasks –

### *An investigation of the iBT-ed journeys of Chinese test takers*

*No activity in educational assessment raises more instructional, ethical, and validity issues than preparation for large-scale, high-stakes tests.*

(Crocker, 2006, p. 115)

There has been a substantial increase in the number of people taking TOEFL iBT worldwide in the recent years. There was a 19% annual increase of Chinese test takers in 2011; and further 32% annual growth in 2012. According to a recent ETS publication (Liu, 2014), Chinese test takers represent about 20% of the TOEFL iBT population. How are Chinese test takers preparing for TOEFL iBT and what are the effects of the preparation practices on their test performance are two central questions of theoretical and practical significance. Test preparation is a contentious issue (Anastasi, 1981; Messick, 1982; Powers, 2012), often debated around effectiveness of preparation on test performance, effect on validity of test scores, equity and fairness of access to opportunity, and impact on learning and teaching more generally. Aiming to achieve positive social and instructional impact (e.g., Cole & Zieky, 2001; Messick, 1980, 1989) of educational assessment products in general, and washback and validity arguments of language tests specifically (e.g., Bailey, 1999; Messick, 1996), ETS has conducted a number of studies on the effects of preparation of high-stakes tests such as GRE, SAT (e.g., Alderman & Powers, 1980; Powers, 1985a, 1985b, 1987, 1993; Powers & Rock, 1999) and TOEFL iBT (e.g., Ling, Powers & Adler, 2014; Liu, 2014). The longitudinal study by Wall and Horák (2006, 2008, 2011) on the washback of the new TOEFL in Central and Eastern Europe exemplifies how a new test can bring about changes in teachers' awareness of and attitudes towards the test and their use of course-books and other test preparation resources.

Since Wall and Horák's study (2006, 2008, 2011), resources for preparing for TOEFL iBT have increased substantially, as the test has been in operation for nearly a decade now. However, relative to resources available for preparing for TOEFL iBT listening, reading and writing tests, the resources and opportunities for developing speaking skills are more limited, especially in

contexts of use such as China. Furthermore, as Chapelle (2008) argued that “as computer-assisted language assessment has become a reality, test takers have needed to *reorient* [our emphasis] their test preparation practices to help them prepare for new test items” (p.127). The delivery of the TOEFL iBT speaking tasks via computers calls for new test preparation practices different from the preparation for tests involving face-to-face interviews (e.g., IELTS speaking test). The TOEFL iBT integrated speaking tasks which require listening/reading skills may present additional challenges and complexities in test taking (Barkaoui, Brooks, Swain, & Lapkin, 2013) and test preparation practices.

As Ross (2008) noted in his editorial of the special issue on English language testing in Asia: “test preparation in some Asian nations has become a massive enterprise which can exert considerable influence against assessment modernization when it comes into conflict with the vested interests of the cram school industry” (p.7). China’s strong tradition of competitive examinations (Martin, 1870) fortifies its booming cram school industry for international English language tests such as TOEFL iBT and IELTS. In order to prepare for the TOEFL speaking tasks, many Chinese test takers choose to attend intensive preparation courses offered by **commercial test preparation schools** – some are big and operating nationwide in an industrial scale, some are small independent businesses or language centres of public universities (hereafter “preparation schools” for all these different types of test preparation operations). Some test takers **may not** attend special preparation courses for various reasons. This could present an equity issue between test takers with different preparation experiences and pose, potentially, threats to test validity, as well as raising concerns about ethical approaches to test preparations (Crocker, 2003; Mehrens & Kaminski, 1989; Popham, 1991). Popham (1991: 13) argued that “No test-preparation practices should increase students’ test scores without simultaneously increasing student mastery of the content domain tested”. Mehrens and Kaminsky (1989) observed that the higher the stakes of the test, the stronger the urge to engage in special test preparation practices. Therefore, they called, as did Popham (1991) and Crocker (2003), for ethical approaches to test preparation, and moral action of different stakeholders, especially teachers and test takers, to ensure test validity and fairness. Not only unethical test preparation but also the variations in ethical preparation could constitute sources of construct-irrelevant variance (Haladyna & Downing, 2004).

To the best of our knowledge, there is no systematic research that has investigated the experiences of Chinese students preparing for TOEFL iBT speaking tasks. What **preparations** (e.g., approaches, practices and materials) do test takers and teachers **do**? **How useful is test preparation**, as perceived by test takers and teachers? To what extent are the perceived effects and effectiveness of test preparation also evidenced in test takers' actual performance? The present study aims to address these questions.

## **Literature Review**

### ***TOEFL iBT Speaking***

The two ETS reports, Douglas (1997) and Butler, Eignor, Jones, McNamara and Soumi (2000), have been instrumental in developing the TOEFL iBT speaking section. The speaking section measures test takers' ability to speak in English effectively in educational environments, both in and outside classroom. It includes six tasks: two independent tasks to express an opinion on topics familiar to test takers; four integrated tasks to speak based on what is read and listened to (Educational Testing Service, 2008). Not only is the entire section new for TOEFL – there was no speaking section on either the paper-based or computer-based TOEFL – but it also represents a significant development and innovation in assessing speaking ability, in its use of integrated tasks (e.g., campus situation and academic course topics) to mirror target language use domains (Bachman & Palmer, 1996). The use of integrated tasks makes the TOEFL iBT speaking tests different from other international English language tests (e.g., IELTS), and may also require or reorient test takers towards new test preparation strategies. Specifically, such integrated speaking tasks require test takers to synthesize and summarize information presented in reading (Yu, 2008, 2009, 2013a, 2013b) and listening materials (Frost, Elder, & Wigglesworth, 2012; Kintsch & Kozminsky, 1977; Lehrer, 1994; Rost, 1994), thus, they represent different constructs from the independent speaking tasks, and as a consequence, the preparation for them is likely to have to be different too. However, this is still very much an assumption; we do not know how the integrated and the independent speaking tasks are being prepared, since to the best of our knowledge there is no empirical data available yet (see also Qian, 2009).

In the *TOEFL Speaking framework*, Butler *et al.* (2000: 23) anticipated that:

*... the introduction of an oral component will have a positive washback effect on the ESL teaching community. By using constructed-response items, which are less likely to be coachable, in the TOEFL 2000 speaking component, we will encourage students to learn to communicate orally – not to learn a skill simply to do well on a test.*

As long as there is an urge to engage in test preparation practices due to the high stakes of a test, special test preparation classes exist anyway, irrespective of whether a test is coachable or not, or whether effects of preparation would be substantive or not. The extent to which the TOEFL iBT Speaking test is coachable, or is sensitive to short-term coaching (Linn, 1990), remains to be evidenced. Similarly, it would be interesting to investigate to what extent the TOEFL iBT Speaking tasks encourage test takers to learn to communicate orally rather than to simply learn a skill to do well on the test.

### ***Test preparation***

#### **Defining test preparation**

The encroaching power of examinations, as Latham (1877) argued, can lead to the special kind of preparation or cramming for examinations. Messick (1982) gave a broadly inclusive definition of coaching or test preparation as: “any intervention procedure specifically undertaken to improve test scores, whether by improving the skills measured by the test or by improving the skills for taking the test, or both” (p.70). In his words, test preparation “may fall anywhere in the broad range bounded by the two extremes of practice and instruction, embracing any combination of test familiarization, drill-and-practice with feedback, motivational enhancement, training in strategies for specific item formats and for general test taking (including advice on pacing, guessing, and managing test anxiety), subject-matter tuition and review, and skill-development exercises” (p.70). Preparation for educational tests is “any procedure specifically oriented toward the improvement of test scores as distinct from nontest-specific learning experiences and cognitive growth” (p.70).

Miyasaka (2000) emphasized five aspects of test preparation: relationship between curriculum and test content, assessment approaches and test formats, test-taking strategies, timing of test preparation and student motivation. Smith (1991), drawing mainly on qualitative interviews with teachers, school administrators, testing experts and school critics, identified eight different

meanings of preparation for external, mandated, high stakes achievement tests in USA elementary schools, with reference to the **micro-politics** of test preparation: (i) ordinary curriculum with no special preparation, (ii) teaching test-taking skills, (iii) trying to exhort pupils to do their best, (iv) teaching content known to be covered by the test, (v) teaching to the test in format and content, (vi) stress inoculation, (vii) practising test or parallel test items, and (viii) cheating during the test. However, this is only one side of the coin, as the eight aspects of test preparation are mainly constructed from the perspectives of stakeholders other than students themselves. The reverse side of test preparation, i.e., students, their views and interpretations of the processes, the meanings and the usefulness of test preparation need to be taken into account in order to draw **a fuller picture** of test preparation practices (see also Hamp-Lyons 1997: 229).

From students' viewpoint, Van Etten, Freebern and Pressley (1997) reported a complex set of beliefs that college students had about examinations: (i) motivation to study for examinations, (ii) strategies for examination preparation, (iii) affect for examination preparation, and (iv) effects of external factors on study, e.g., instructors, previous examination experiences, social environment, physical environment, and content to be studied (see Van Etten et al., p.201 for detailed descriptors of these aspects). The aspects of test preparation identified by Van Etten et al. (1997), Miyasaka (2000), Smith (1991) and Messick (1982), together present an important theoretical framework for the present research to investigate the process and effect of dedicated preparations for TOEFL iBT speaking test. In the following sections, **we review the research studies on the process and effect** of test preparation with specific reference to IELTS and TOEFL, and the role of materials **in test** preparation.

### **Preparing for IELTS and TOEFL: Process and effect**

In the field of language testing, test preparation is often conceptualized and investigated in relation to ethics, fairness and washback. Research on washback of high-stakes language tests on day-to-day classroom teaching and learning as part of “normal” school curriculum is well documented (e.g., Alderson & Wall, 1993; Bailey, 1996; Cheng, Watanabe, & Curtis, 2004; Rea-Dickins & Scott, 2007; Spratt, 2005; Wall, 1996; Wall & Alderson, 1993; just to name a few). However, unlike day-to-day teaching and learning as part of school curriculum, special test **preparation programmes** are often external or additional to normal school curriculum and are dedicated specifically to enhancing test performance as its main, if not sole, **purpose**. **Here** we

focus mainly on the research studies of special preparation courses for the two major international English language tests TOEFL and IELTS<sup>1</sup>. From the perspectives of different stakeholders (Rea-Dickins, 1997), research on special coaching programmes for international English language tests has examined (i) how students prepare or are being prepared for TOEFL (e.g., Alderson & Hamp-Lyons, 1996; Hamp-Lyons, 1998, 1999; Matoush & Fu, 2012; Wall & Horák, 2006, 2008, 2011) and IELTS (Badger & Yan, 2012; Everett & Colman, 2003; Gibson & Swan, 2008; A. Green, 2006; Hayes & Read, 2004; Mickan & Motteram, 2008; Read & Hayes, 2003; Saville & Hawkey, 2004), and (ii) the effects of intensive preparations on test performance or score gains or improvements in language proficiency (e.g., Brown, 1998; Elder & O'Loughlin, 2003; Gan, 2009; A. Green, 2005, 2007; Ling, Powers, & Adler, 2014) and the differential effects of various test preparation strategies on test performance (e.g., Issitt, 2008; Liu, 2014). The majority of such studies on the effects of special test preparation programmes have tended to investigate the overall improvement of test performance covering all language skills, with a smaller number of studies focusing on a specific language skill.

A number of studies have investigated the features of IELTS test preparation and effects on performance. Saville and Hawkey (2004) presented an overview of the IELTS Impact Study and its sub-projects, and reported in particular the development and validation of the instruments to evaluate test preparation course books. Everett and Colman (2003) evaluated the appropriateness of the content, organisation, learning approaches and presentation of the listening and reading components of six commercial course books widely used for IELTS preparation at three Australian language centres. They argued that course books purporting to prepare students for IELTS should include more texts and tasks that can contribute positively to the social and academic acculturation of students, in addition to simulating practice tests. Read and Hayes (2003) and Hayes and Read (2004) surveyed the provision of IELTS preparation courses in New Zealand and compared two IELTS preparation courses at university language centres in Auckland – one almost entirely IELTS-focused and the other as an elective within a General English programme. They reported a number of substantial differences in the focus and delivery

---

<sup>1</sup> We are aware that there are a number of studies on how test takers prepare for other English language tests such as FCE and College English Test. However, due to the different nature of the stakes of the tests, we consider the findings of research on TOEFL and IELTS-Academic more comparable in many ways and therefore have more direct relevance to the focus and methodology of the present study. Having said this, we are by no means ignoring the findings of research studies on intensive preparation for other English language tests; some of them (e.g., Bachman, Davidson, Ryan, & Choi, 1995; Farnsworth, 2013; Lumley & Stoneman, 2000; Xie, 2013; Xie & Andrews, 2013) are referred to in this report where appropriate.

of the two courses. Similarly, Green (2006) compared IELTS writing preparation courses against EAP writing courses, and found that the IELTS-focused test preparation strategies were not driven primarily by students' expectations of the courses. Mickan and Motteram's (2008) observational study of an intensive eight-week IELTS preparation programme evidenced an eclectic instructional approach which constituted a complex process of socialisation into test-taking behaviours governed by the priorities of the test tasks. The socialisation included essentially the process of students' familiarising, practising and rehearsing test tasks with teachers modelling and scaffolding exemplar texts and giving practical hints and strategies for doing the tasks. Gibson and Swan (2008) examined how the sociolinguistic backgrounds of the non-native English teachers of IELTS preparation courses in Malaysia might affect their understanding of the test construct and their delivery of the preparation courses. Like Mickan and Motteram (2008), Badger and Yan (2012) examined the teaching methodology used in IELTS preparation courses in China, in terms of their pedagogical orientation, instructional content and presentation, language activities, roles of teachers and learners, teaching materials, and assessment methods. They found that the classes were largely communicative but teacher-centred with more frequent use of Chinese than English as the medium of instruction.

The studies reviewed in the paragraph above focused mainly on the process of IELTS preparation courses. Below we report briefly the main findings of the studies on the effects of intensive preparation on test performance. Brown (1998) and Green (2007) compared not only IELTS-focused and EAP-focused programmes but also the effects of these programmes in improving IELTS writing test scores. Brown (1998) reported an average gain of one band score for IELTS Academic writing over a ten-week course of instruction. However, Green's study (2007) showed no clear advantage of IELTS-focused preparation in improving IELTS writing scores. Elder and O'Loughlin (2003) found that on average students gained about half a band score overall during a nine-month test preparation programme, with the greatest improvement in listening and the least in reading. However, as Green (2005) heeded, length of test preparation course is not as successful for predicting test takers' score gains as their initial English language ability at the point of starting preparation course. Elder and O'Loughlin (2003) lent further support to this – they found that student's educational qualification was the best predictor of score gains. From a slightly different angle, Gan (2009) attributed the non-significant difference



in IELTS scores between students who had taken preparation courses and those who had not to the narrowing gap between students' overall English language proficiency after taking both general English courses and IELTS-focused preparation courses. However, it should be noted that Gan (2009) did not provide information about the IELTS-focused preparation courses, e.g., who taught the preparation courses, what and how test preparation materials were used.

Alderson and Hamp-Lyons (1996) observed both TOEFL preparation classes and non-TOEFL preparation classes of two teachers in North America for one week's duration, and conducted interviews with teachers and students (including those from other test preparation institutes). Students were asked to suggest how they think TOEFL preparation courses should be taught, in comparison to what they had experienced. The classroom observation data showed that there were "substantial differences between TOEFL and non-TOEFL classes" (p.289), and the differences between the two teachers were "at least as great as the differences between TOEFL and non-TOEFL classes" (p.290). Hamp-Lyons (1998) critiqued five textbooks selected at random from those on the market targeting specifically TOEFL test takers (see also Hilke & Wadden, 1997; Wadden & Hilke, 1999). She found that the textbooks emphasised primarily two skills: "test-taking strategies and mastery of language structures, lexis, and discourse semantics that have been observed on previous TOEFLs" (p.332), and that the textbooks did "little with task types or item formats other than those predicted to occur on the TOEFL on the basis of analysis of past forms" (p.334). To understand the impact of the introduction of the new TOEFL, Wall and Horák (2006, 2008, 2011) conducted a longitudinal, qualitative study in Central and Eastern Europe to track a small number of teachers on their awareness of and attitudes towards the test and their use of commercial textbooks and other test preparation resources. Matoush and Fu (2012) compared their own experiences in teaching TOEFL iBT in China as native and non-native speakers of English respectively, juggling with customers' (especially **students'**) and employers' expectations for short-term testwiseness with their own understandings of long-term language and literacy goals for academic study. These studies on TOEFL preparation, though based on a small sample of participants, evidenced the complex, almost idiosyncratic, nature of test preparation courses which were influenced by teachers' personal and professional characteristics among many other factors. None of these studies, however, looked at the effect of intensive preparation on performance.

Studies on the effects of dedicated TOEFL iBT preparation courses (e.g., Ling, et al., 2014) and test takers' preparation strategies are emerging (e.g., Liu, 2014). Ling et al. (2014: 14) found that students in one participating school in China achieved, after taking the 9-month intensive preparation courses, “moderate to substantial improvement” in English skills and TOEFL iBT test scores especially on listening and reading tests. Liu (2014: 11) conducted a much larger survey with Chinese test takers. It was found that whether attending a coaching school or not “has a fairly weak relationship with the reading and listening skills assessed by the TOEFL iBT and has almost no relationship with writing and speaking”. Nevertheless, she found that different types of preparation strategies (TOEFL-focused vs. more general language learning strategies) had differential effects on test performances (in terms of both TOEFL iBT total score, and the sub-scores of reading, listening, writing and speaking). For example, practising TOEFL-like simulation tests or released items and memorizing vocabulary had the largest effect on test takers' total TOEFL scores. “Practice spoken English using templates (e.g., use common transitional phrases, use common argument structure)” and “improve fluency in speaking” were the two best predictors of Speaking scores. Ling et al. (2014) and Liu (2014) seemed to suggest Speaking was much less coachable than other skills in TOEFL iBT tests. However, it must be pointed out that none of the TOEFL studies investigated Speaking test preparation specifically, which is the focus of the present study.

### **Role of test preparation materials**

The majority of the studies reviewed above point to the fact that test preparation materials (course books as well as other resources) can play a mediating role in the process as well as the effect of test preparation. Test preparation materials are an integral and most tangible aspect of test preparation as they define and determine to a large extent strategies and approaches of test preparation and consequently affect the effect and efficiency of test preparation. Learning materials are the concrete expression of programme purpose and objectives, and frame, in the same activity space, the contributions of teachers and learners (Johnson, 1989; Kiely & Rea-Dickins, 2005; Tomlinson, 2003, 2012). Hamp-Lyons (1998: 134) argued that if the content and design of test preparation materials can support teachers to help their learners to increase their language proficiency, and at the same time the test preparation teachers see their principal task as helping their learners to increase their language proficiency and consciously choose appropriate content and methods there might be beneficial washback. Wall and Horák (2006,

2008, 2011) has demonstrated the vital importance to consider the teacher's role in using test preparation materials in response to innovations such as a new test format. However, equally important are the view and approaches that students would take. Lumley and Stoneman (2000) found that teachers "clearly saw the potential of the materials as a teaching package, ... including but extending beyond test preparation", while students "were above all concerned with familiarising themselves with the format of the test, ... demonstrated relatively little interest in the idea of using test preparation as an opportunity for language learning" (p.70).

### ***Research gaps and the way forward***

Studies reviewed above suggest we have accumulated a significant amount of research evidence on how high-stakes language tests are prepared for in different locations, with regard to the process (strategies, approaches and materials) and effect of test preparation programmes. However, there are at least three notable research gaps. Firstly, although the theoretical frameworks of test preparation proposed by Messick (1982), Miyasaka (2000), Smith (1991) and Van Etten et al. (1997) are usefully operationalizable for systematic investigations of various aspects of test preparation, there is clear evidence of differences in their interpretations as to what constitutes the most important aspects of test preparation. Secondly, the studies on IELTS and TOEFL test preparation courses tended to focus on a limited number of aspects of test preparation, thus failing to treat test preparation as a dynamic system. For example, Hamp-Lyons (1998), Everett and Colman (2003), Saville and Hawkey (2004), and Wall and Horák (2006, 2008, 2011) focused on course books. The research studies comparing the features of test preparation and non-test-preparation courses tended to rely on classroom observations to unpack the differences in their instructional approaches or teaching methodology. Other studies tried to examine how characteristics and qualifications of teachers and students affected their test preparation strategies. Collectively these studies make important contributions to our understanding of test preparation; however, individually these studies were quite fragmented in terms of their research focus and approaches. Thirdly, there is a dearth of research focusing on intensive preparation for a Speaking test. Even though some of the studies did touch upon preparation for IELTS Speaking test, any such findings of these studies are not really applicable to the TOEFL iBT Speaking test, since the underlying construct and the formats of the two tests are quite different: the TOEFL iBT Speaking test includes both independent and integrated tasks and are computer mediated, while the IELTS Speaking test includes monologues and interviews

with human examiners. We also argue that test takers' different preparation experiences (whether and what kind of coaching programmes they attend) might present some equity, validity and ethical issues. Understanding how test takers prepare for a test contributes to establishing the validity argument of the test, since test preparation constitutes sources of construct-irrelevant variance (Haladyna & Downing, 2004).

As Hamp-Lyons (1997: 299) noted: "Many more studies are needed of students' views and their accounts of the effects on their lives of test preparation, test-taking and the scores they have received on tests". Alderson (2004) commented "there have been fewer studies of what students think, what their test preparation strategies are and why they do what they do, but we are starting to get insights" (p. 2), but "so little of teachers' motives for teaching test-preparation lessons the way they do is ever addressed critically in the literature" (p.5). We therefore stress the importance of understanding the process, the meanings and the usefulness and effects of test preparation from the perspectives of the two key stakeholder groups (Rea-Dickins 1997) – test takers and teachers – from a holistic and contextualized approach. In this approach, we view the local context not only as the central player for shaping the current test preparation market and practice but also as the key to understanding and interpreting our research findings. China is a highly examination-oriented society (Yu & Jin, 2014, 2015), and its economic boom in the **last** fifteen years has further enhanced the test preparation market. Different preparation schools and their teachers may compete and contend to seek to be the critical reality definers, for relative influence and control over resources, reputations, respect and most importantly their share in the highly competitive and lucrative test preparation market in China.

## **Research Questions**

This study aimed to address the issues identified above through four research questions (RQ).

**RQ 1: How do Chinese students prepare for TOEFL iBT speaking tasks?**

**RQ 2: In what ways do students who attend intensive preparation courses and those who do not differ in their test preparation?**

**RQ 3: In what ways is test preparation useful from the perspectives of students and teachers?**

**RQ 4: What are the relationships between test preparation and student’ actual performance in TOEFL iBT speaking test?**

## **Method**

### ***Participants: test takers and teachers***

This project collected data from TOEFL iBT test takers and teachers in four major cities in China – Hangzhou, Nanjing, Shanghai and Beijing. At the time of our data collection, all student participants were intensively preparing for TOEFL iBT and aiming to take the test within half a year or so, or had just taken the test in the past four months. Their preparation for TOEFL iBT speaking test can be via commercial test preparation schools or on their own (i.e., not attending intensive preparation courses). The student participants who did not attend special preparation courses were recruited from some universities in the four cities with which we had connections, as well as via our online student questionnaire survey – the recruitment advertisement with the link to our online student questionnaire was posted in a number of popular test preparation websites in China. The teacher participants were recruited from test preparation schools. Before taking part in this project, all the participants signed the consent form.

### ***Instruments and procedure***

#### **Understanding the context of TOEFL iBT test preparation**

In order to gain an overview of what was going on in TOEFL iBT Speaking test preparation market, we visited a number of websites targeting specifically Chinese test takers, including those of test preparation schools, open discussion forums and book stores, for example:

- [www.5libt.cn](http://www.5libt.cn)
- [www.taisha.org/test/toefl](http://www.taisha.org/test/toefl)
- [www.igo99.cn/toefl/](http://www.igo99.cn/toefl/)
- [www.onlytoya.com/](http://www.onlytoya.com/)
- [www.gter.net/list-5-1.html](http://www.gter.net/list-5-1.html)
- [www.manfen.com](http://www.manfen.com)
- [www.etest.net.cn](http://www.etest.net.cn)
- [toefl.xdf.cn/](http://toefl.xdf.cn/) ([toefl.koolearn.com/](http://toefl.koolearn.com/))
- [www.xhd.cn](http://www.xhd.cn) ([www.xhd.org](http://www.xhd.org))
- [www.longr.com](http://www.longr.com)
- [en.eol.cn/zt/201211/toefl\\_speak/](http://en.eol.cn/zt/201211/toefl_speak/)
- [toefl.eol.cn](http://toefl.eol.cn)

- [www.manfen.net](http://www.manfen.net)
- [www.xiaomajiaoyu.com/toefl/](http://www.xiaomajiaoyu.com/toefl/)
- [www.91toefl.com/](http://www.91toefl.com/)
- [www.ntoefl.com.cn/](http://www.ntoefl.com.cn/)
- [www.sisutoefl.com/](http://www.sisutoefl.com/)
- [www.exam8.com/english/TOEFL/](http://www.exam8.com/english/TOEFL/)
- [www.liuxue86.com/toefl/](http://www.liuxue86.com/toefl/)

All sorts of information are available from these websites, from advertisements of test preparation courses, freely downloadable course books and computer software, to the so-called “Ji-Jing” discussion boards (where test takers shared their experience of taking TOEFL iBT or the speaking tasks they remembered). It was an essential first step for us to understand the context of our research to inform the design of data **collection tools**.

Furthermore, we conducted seven initial focus-group discussions, in Hangzhou, Nanjing and Shanghai, with test takers and teachers of preparation schools as well as some university students who did not attend any dedicated TOEFL iBT preparation course. In order to maximize the range of possible comments the participants would make during the focus-group discussions and to produce a framework that would capture as wide a range of perceptions and practices as possible about TOEFL iBT speaking test preparations, we asked the participants some general questions such as: why, when, where, and how do you prepare for or teach TOEFL iBT Speaking test? What materials do you use to prepare for or teach TOEFL iBT Speaking test? What are your suggestions for us in designing the questionnaires for the research project?

### **Test taker and teacher questionnaires**

The seven initial focus-group discussions were audio-recorded, transcribed and coded to assist the development of the test taker and teacher questionnaires. The design of the questionnaires was also informed by (i) the information we collected from the websites described above and (ii) the research literature on test **preparation, in particular, those learning, instructional and affective aspects of test preparation as identified by Messick (1982), Miyasaka (2000), Smith (1991) and Van Etten et al. (1997). These aspects of test preparation were also investigated in a number of research studies on IELTS and TOEFL courses.** The two questionnaires aimed to capture an overview of the strategies and materials used by test takers and teachers preparing for TOEFL iBT speaking tasks. Drafts of both questionnaires were sent to the teachers who participated in the initial focus-group discussions for comments and suggestions. Five teachers responded with

suggestions for improving the design and content of the questionnaires. After further revisions, the student questionnaire was then piloted in Hangzhou, Nanjing and Shanghai.

The final version of the student questionnaire (see Appendix 1) consisted of three sections and covered the following topics.

#### Section 1: Demographic data

- name, gender, age, current city of residence, mobile phone number, email address,
- education level, university degree and specialism,
- English language proficiency as demonstrated in other standardized tests such as College English Test, Test for English Majors, and IELTS, number of years learning English.

#### Section 2: General test taking and preparation experience

- experience of and plan for taking TOEFL iBT, and TOEFL iBT scores achieved (for those who had taken the test before our data collection),
- purpose of taking the test,
- amount of time spent preparing for the independent and integrated Speaking tasks in comparison to Reading, Listening and Writing tasks,
- difficulty level of the independent and integrated speaking tasks,
- frequency and usefulness of 17 types of test preparation tasks (e.g., “read aloud”, “summarize orally”, “memorise model essays”, “do mock tests”, “study rating criteria”, and “learn about TOEFL iBT Speaking related topics”),
- frequency and usefulness of ten widely available test preparation course books,
- five most frequently used test preparation materials, websites, and test taking strategies,
- whether or not taking intensive test preparation lessons; if not, why.

#### Section 3: Experience at test preparation school

- General information about the programme: name of the institution and programme, location, number of hours for the whole programme and for the Speaking test, time of speaking lessons, ratio of time spent on independent and integrated speaking tasks, number of students in speaking lessons, medium of instruction,
- purposes, expectations and other reasons for taking preparation course,

- frequency and usefulness of 13 tasks that teachers do during the lessons (e.g., “teachers lecturing on test taking strategies”, “teachers providing sample answers”, “doing mock tests”, “teachers explaining scoring rubrics”, “teachers correcting student grammatical mistakes”, “teachers correcting student pronunciation and intonation”, “teachers assigning homework”),
- usefulness of preparation lessons for improving the chance of getting a higher score in TOEFL iBT Speaking.

The Chinese version of student questionnaire (available from the first author upon request) was administered outside normal lesson time for all students, at different point of time in different test preparation schools. In order to reach those test takers who were preparing for the TOEFL iBT Speaking test on their own, we also developed an equivalent online version of the questionnaire. However, this proved to be not as successful as we had hoped, although we took every effort to recruit participants who were preparing TOEFL iBT on their own. In total, we collected 1,514 valid questionnaires from students, but only 70 said they were not attending any special test preparation course. We therefore decided that we had to drop RQ2 (In what ways do students who attend intensive preparation courses and those who do not differ in their test preparation?), as the findings from comparing 70 participants with 1444 are not that meaningful, if not misleading.

The teacher questionnaire (see Appendix 2) was designed along the same line with the student questionnaire; it consisted of two sections and covered the following topics, most of which had corresponding items in the student questionnaire (some in slightly different wording). The Chinese version of the questionnaire (available from the first author upon request) was administered at a time convenient for the teacher. In total we collected 23 valid questionnaires from teachers.

#### Section 1: Demographic data

- name, gender, age, mobile phone number and email address,
- university degree and specialism, employer (i.e., test preparation school/centre), employment status (part-time or full-time), experience in teaching English, TOEFL



preparation classes (including old and new TOEFL) and TOEFL iBT Speaking classes, pre- and in-service training received for teaching TOEFL iBT Speaking classes.

## Section 2: Experience in teaching TOEFL iBT Speaking

- TOEFL iBT Speaking test preparation courses taught (number of hours and students),
- medium of instruction,
- ratio of time spent on independent and integrated tasks,
- frequency and usefulness of 13 activities that teachers do during the lessons,
- frequency and usefulness of 35 test preparation tasks that teachers organize students to do in the lessons (These were collapsed into 17 tasks in the student questionnaire, Section 2),
- frequency and usefulness of twenty-three widely available test preparation course books (Ten of them were listed in the student questionnaire),
- five recommended websites, and reasons for recommendation,
- five recommended test taking strategies, and reasons for recommendation,
- special internal textbooks and other materials used,
- group or individual preparation for the lessons,
- students' purposes, expectations and other reasons for taking preparation course, value of taking preparation course,
- personal experience of taking TOEFL iBT: test score and purpose of taking the test,
- self-assessment of speaking abilities, and knowledge about speaking tests in general and TOEFL iBT speaking test and rating criteria specifically,
- demands of teaching TOEFL iBT Speaking classes in comparison with other speaking classes.

## Interviews and classroom observations

Upon completion of the **questionnaires**, we conducted a number of **small group and** one-to-one interviews, depending on the participants' availability. Wherever possible, the students and teachers were interviewed separately. The questions we asked **at the interviews followed** broadly the questions we asked at the questionnaires (see Appendix 3 for the procedure and questions we asked at student interviews; Appendix 4 for teacher interviews), **which** further explore their experiences in preparing for TOEFL iBT Speaking test. The interviews were conducted largely in Chinese, with occasional use of English where appropriate, and were all audio-recorded (except

for one teacher interview). In total, we interviewed 53 students and 33 teachers. All interviews were transcribed verbatim for analysis.

A small number of teachers agreed for us to observe and audio-record their lessons. In addition to recording the lessons, we used a specifically designed classroom observation form (Appendix 5) to capture in real-time the focus of the episodes of the lessons, e.g., what resources teachers used, what advice and strategies teachers gave for independent and integrated tasks, what specific speaking opportunities or tasks that students were given in class, and how much time was spent on different classroom activities. The classroom data would provide a valuable validity check on the representation of test preparation activities emerging from the interviews, focus-group discussions and questionnaire data. In total, we observed 7.5 hours of TOEFL iBT Speaking lessons, which were transcribed verbatim for analysis. The qualitative data of focus-group discussions and one-to-one interviews and classroom observations are supplementary to aid our interpretation of the quantitative questionnaire data.

### **TOEFL iBT test scores**

Unlike Liu (2014), we were not able to obtain our participants' official TOEFL iBT scores directly from ETS data warehouse. Instead, we phoned from Bristol 1221 students who provided a valid telephone number in the questionnaire. At the same time, we also sent a personal email to every student (n=1337) who had provided an email address in the questionnaire, asking them to tell us their TOEFL test scores via email or an online questionnaire we set up for collecting test scores. However, not every student was willing to share their test scores with us, or had taken the test by our cut-off time (October 2012). Eventually, we managed to obtain test scores from 293 students. However, not all of them provided the full set of sub-scores; some provided total scores or Speaking scores only. The small sample size for TOEFL iBT scores limited our ability to fully address RQ4 (What are the relationships between test preparation and students' actual performance in TOEFL iBT speaking test?)

### ***Data analysis***

Descriptive analyses were conducted on the student participants' background variables (e.g., age, gender, education level/status, experience in taking other international English language tests, experience in learning English as a foreign language) and TOEFL iBT-related variables (e.g.,

previous experience in taking TOEFL iBT, purpose of taking/preparing for TOEFL iBT, frequency and usefulness of various test preparation strategies and practices, expectations of taking dedicated intensive test preparation courses). Similar descriptive analyses were conducted on the teacher participants' background variables as well as the variables about their experience in teaching and taking TOEFL iBT Speaking. We also conducted factor analysis on the overall data of test preparation (55 items/questions covering a range of test preparation activities, course-books, and reasons and expectations for taking preparation courses), as well as separate factor analysis on the data of students' reasons and expectations for taking preparation courses. A key decision in factor analysis is what rotation method to use, which is both a mathematical and philosophical question about the underlying construct under investigation. "Many have argued that correlated factors are much more reasonable to assume in most cases..., and therefore oblique rotations are quite reasonable." (Stevens, 2002: 392). Pedhazur and Schmelkin (Pedhazur & Schmelkin, 1991: 615, cited in Stevens 2002: 392) argued that:

From the perspective of construct validation, the decision whether to rotate factors orthogonally or obliquely reflects one's conception regarding the structure of the construct under consideration. It boils down to the question: Are aspects of a postulated multidimensional construct intercorrelated? The answer to this question is relegated to the status of an assumption when an orthogonal rotation is employed... The preferred course of action is, in our opinion, to rotate both orthogonally and obliquely. When, on the basis of the latter, it is concluded that the correlation among the factors are negligible, the interpretation of the simpler orthogonal solution becomes tenable.

We first rotated the data obliquely (direct oblimin) with the desired number of factors and looked at the correlations among the factors; it was noted that correlations among some factors were below 0.30 (i.e., less than 10% overlap in variance among factors). We also ran the data orthogonally (varimax). For ease of interpretation of the factors, we report the findings based on varimax rotation method.

Due to the small number of test scores we collected from students, we ran simple regression

analysis, instead of multi-level modelling as we had proposed, to identify the relationship between students' use of different types of test preparation strategies and their actual test results. The supplementary data of recorded interviews and lessons were analysed to provide further explanatory power for the interpretation of the questionnaire data. Any anomaly identified from the interviews and classroom observations (i.e., difference between the qualitative data and the large-scale quantitative questionnaire data) would provide further insights into special cases of test preparation.

## Results

Drawing on and integrating the two quantitative datasets (students and teachers questionnaires), we report the characteristics of the participants, the multiple aspects of test preparation, e.g., students' time management, students' and teachers' perceptions of the difficulty level of the independent and integrated speaking tasks, the frequency of use and the perceived usefulness of a wide range of test preparation activities, features of course-books and other test preparation materials, reasons and expectations for taking test preparation courses, features of test preparation courses, and learning of test-taking strategies. We also report how students' gender, age and test-taking purposes – three key variables that are often assumed to have some association with people's language learning styles, strategies, motivation and achievement (J. M. Green & Oxford, 1995; Oxford, 1989) – might have affected how they prepare for the speaking test. Based on the results of factor analysis on the student questionnaire data, we propose a five-component framework to capture the complexity and dynamics of test preparation. Finally we report the relationships between test preparation and test performance.

### *Characteristics of student participants: Age, gender and educational qualifications*

We collected 1514 valid student questionnaires: 1464 were paper-based and 50 online. The vast majority (1508) were collected from the four targeted cities. Only 70 students were preparing TOEFL iBT on their own. The tuition fees of the vast majority of the students (91.2%) attending preparation courses were paid by their parents. The students were equally distributed in terms of gender; and nearly 90% of them were between 16-24 years of age. About 1/3 the students were studying in secondary schools, and nearly 60% studying in university or college. There was some

small discrepancy between the number of questionnaires we collected from a city and the number of students who said they lived in that city. This was largely due to some students coming all the way from other, often smaller, cities in order to study at their preferred test preparation school in one of the four big cities. In addition, one company in Shanghai arranged their employees to be trained full-time in a test preparation school in Hangzhou.

The majority of the students (74.4%) have been learning English for 8-12 years (mini.=1, maxi.=34,  $M=10.19$ ,  $SD=3.02$ ). About 10% of the students ( $n=154$ ) had taken TOEFL iBT; and their TOEFL iBT scores ranged from 33 to 115 ( $M=85.69$ ,  $SD=18.78$ ,  $n=110$ ).

Over 91% of the students would take TOEFL iBT for academic degree study abroad, with 43.1% for Master, 36.8% for Bachelor, 6.5% for inter-university student exchange, and 4.9% for Doctoral degree programmes. The “other reasons” include mainly “studying in secondary school in the USA”, and “required by my company for its internal selection of candidates for training overseas”. Like those participants in Liu (2014), the vast majority of the participants in the present project would take TOEFL iBT for admission to academic programmes. Liu (2014) reported that 88% of the respondents ( $N=14,593$ ) to her online questionnaire sought admission to college (19%) and graduate school (69%). The present project had a higher percentage of participants (36.8%) who aimed to study for their first degree than Liu (2014).

### ***Preparing for TOEFL iBT Speaking tasks***

Test preparation is a complex endeavour, involving a host of intertwined strategic decision-making such as managing and allocating time for different tasks and skills, and prioritizing different test preparation activities, course-books, and test-taking strategies, in order to maximize the effects of intensive preparation on test performance.

### **Time management and commitment**

At the time of our data collection 67.2% of the students planned to take TOEFL iBT within half a year, and 61.6% of students had already registered for the test. Even for those who had not registered for the test, over 50% of them said they planned to take the test within three months. We asked the students how long it was between their registered test date and the date when they decided to attend intensive test preparation course, 77% of them said it was under 6 months, and

14.6% between 6 and 12 months, only 8.4% were over 12 months. It is evident that the majority of the students were taking intensive preparation courses quite close to their planned/registered test date, seeking some short-term effects to improve their test scores. The majority of the students (81.9%, valid N=1263) took “day” courses, 4.4% took “evening” courses and 13.7% took both day and evening courses. The majority of the students (70.7%, valid N=1258) attended Speaking lessons at both weekdays and weekends. Just under a ¼ attended their lessons during weekdays only.

The students were also asked to rank the amount of time they spent on the four language skills during test preparation (see Table 1). Listening was prioritized by the majority of the participants (61.9%), followed by Reading (17.7%), Speaking (11.6%) and Writing (8.8%). Only 11.6% of the participants considered Speaking as their No.1 priority, 37.1% had Speaking as their No.2 priority, and the remaining 50% considered Speaking as either No.3 or No.4 skill – equally distributed. Overall, it seems that the receptive skills (Listening and Reading) have received more attention than the productive skills (Speaking and Writing) in test preparation.

**Table 1 Rank order of the amount of time spent on the four language skills**

Skills	No.1		No. 2		No. 3		No. 4	
	Freq	Valid %	Freq	Valid %	Freq	Valid %	Freq	Valid %
Listening	862	61.9	288	20.7	172	12.4	71	5.1
Speaking	161	11.6	517	37.1	357	25.7	358	25.7
Reading	247	17.7	368	26.5	388	27.9	388	27.9
Writing	123	8.8	219	15.7	474	34.1	574	41.3
Valid N. Total	1393		1392		1391		1391	

Further analysis of the students’ preference for their No.1 skill was conducted in relation to the student’s gender, age (the 16-18 year olds vs. 19-24 year olds), purpose of taking TOEFL iBT (those aiming to study for their first degree vs. those for their **Master degree**). No significant difference between male and female students in their priority for Listening or Writing was found. However, overall, older students were more likely to prioritize Listening than younger students, with the 19-24 age group ( $M_{19-24\text{years}}=1.60$ ,  $SD=.937$ ,  $n=791$ ) spending significantly more time in Listening than the younger 16-18 age group, **though with a small effect size** ( $M_{16-18\text{years}}=1.93$ ,  $SD=1.142$ ,  $n=441$ ;  $t=5.473$ ,  $df=1230$ ,  $p<.0005$ , **Cohen’s  $d=.325$** ). The younger students, the

16-18 year olds spent more time in preparing for Speaking than the 19-24 year olds did. Students aiming to study for their first degree ( $M_{\text{firstdegree}}=1.92$ ,  $SD=1.128$ ,  $n=491$ ) were more likely to prioritize Speaking than those aiming to study for their Master degree ( $M_{\text{masterdegree}}=1.57$ ,  $SD=.918$ ,  $n=593$ ), and the difference between them was statistically significant ( $t=5.621$ ,  $df=1082$ ,  $p<.0005$ , **Cohen's  $d=.344$** ). As the age groups were roughly in line with the groups of test taking purposes in the sample, it was no surprise that similar findings **in these two comparisons were observed**.

**As TOEFL iBT Speaking test** requires not just speaking skill, we asked the students what percentage of **time they** spent preparing for the Speaking test. It was found that the majority of the students (56.9%) spent 20-40% of their time preparing for the Speaking test (see Table 2). No significant difference was observed between genders, age (16-18 vs. 19-24), **and test-taking purposes (studying for undergraduate vs. Master programmes)**, indicating that 20-40% of time spent on the Speaking test seemed to be the norm across the board.

**Table 2 Percent of time spent on Speaking test**

Percent of time on Speaking test	Freq	Valid %
<20% (1)	244	17.3
>=20%-40% (2)	804	56.9
>=40%-60% (3)	280	19.8
>=60%-80% (4)	62	4.4
>=80%-100% (5)	22	1.6
Valid N.	1412	

On average, the students spent around 14 hours a week in the last month ( $M=13.96$ ,  $SD=13.67$ ) preparing for the Speaking test. The large standard deviation indicated the big variation among the students. Furthermore, we asked the students what ratio of time they spent on independent and integrated Speaking tasks. It was found that over  $\frac{3}{4}$  of the students ( $n=932$ ) spent more time on integrated tasks than independent tasks; 15.7% ( $n=194$ ) equally allocated their time to independent and integrated tasks; and 8.9% ( $n=110$ ) spent more time on independent than integrated tasks. Specifically, 102 students spent 300% more, 453 spent nearly 150% more, and 364 spent nearly 50% more time on integrated than independent speaking tasks. In contrast, a much smaller number of students spent more time on independent than integrated tasks; specifically, only 12 students spent 300% more, 42 spent nearly 150% more, and 50 spent 50% more time on independent than integrated tasks. A further question in the Student Questionnaire

(No. 3.9, See Appendix 1) asked the students the ratio of time their teachers spent on independent and integrated tasks. The data confirmed that teachers also spent a lot more time on integrated than independent speaking tasks during lessons. Nearly 70% of the students thought their teachers spent more time on integrated than independent tasks; 22.5% thought it was equal, and only 7.8% thought more time was spent on independent tasks. This finding was further corroborated with the teacher questionnaire data too (see No.2.4 of the Teacher Questionnaire, Appendix 2). Only one teacher said that she spent more time on independent than integrated tasks; three teachers said equal time was spent on the two types of tasks, and the rest said they spent more time on integrated than independent tasks. According to both student and teacher questionnaire data, the most common ratio of time allocated to preparing for independent and integrated speaking tasks was 30:70. In other words, in most cases, the time spent on integrated speaking tasks was about twice of the time on independent tasks.

### Perception of the difficulty level of the Speaking tasks

The fact that the participants allocated substantially more time to integrated than independent speaking tasks may well reflect the perceived difficulty level of the two types of tasks ( $M_{\text{independent}}=3.39$ ,  $SD=.873$ ;  $M_{\text{integrated}}=3.87$ ,  $SD=.833$ ). As shown in Table 3 below, independent Speaking tasks were considered easier than integrated Speaking tasks.

**Table 3 Difficulty of the Speaking tasks**

Difficulty of Speaking tasks	Independent tasks		Integrated tasks	
	Freq.	Valid %	Freq.	Valid %
very easy (1)	16	1.1	16	1.2
Easy (2)	161	11.5	47	3.4
somewhat difficult (3)	640	45.9	350	25.2
Difficult (4)	422	30.3	669	48.2
very difficult (5)	156	11.2	306	22.0
Valid N	1395		1388	

Further analysis indicated that male students rated the independent tasks as more difficult than female students ( $M_{\text{male}}=3.45$ ,  $SD=.899$ ,  $n=657$ ;  $M_{\text{female}}=3.34$ ,  $SD=.844$ ,  $n=687$ ;  $t=-2.216$ ,  $df=1342$ ,  $p<.0275$ , **Cohen's d=.126**), but not the integrated tasks ( $M_{\text{male}}=3.89$ ,  $SD=.846$ ,  $n=656$ ;  $M_{\text{female}}=3.85$ ,  $SD=.819$ ,  $n=685$ ). The older students considered the integrated tasks significantly more difficult than the younger ones ( $M_{16-18\text{years}}=3.79$ ,  $SD=.894$ ,  $n=440$ ;  $M_{19-24\text{years}}=3.94$ ,  $SD=.796$ ,  $n=790$ ;  $t=-3.021$ ,  $df=1228$ ,  $p<.0035$ , **Cohen's d=0.18**), but not the independent tasks ( $M_{16-18\text{years}}=3.38$ ,  $SD=.909$ ,  $n=441$ ;  $M_{19-24\text{years}}=3.41$ ,  $SD=.844$ ,  $n=792$ ;  $t=-.590$ ,  $df=1231$ ,  $n.s.$ ).



Students taking TOEFL iBT to apply for undergraduate programmes were similar to those applying for Master programmes in their perception on the difficulty level of independent and integrated tasks.

### Frequency and usefulness of different test preparation activities

As shown in Table 4 below, 16 of the 17 activities can be done by the students on their own, and only one activity (“talk to people in English”) has to involve another person. It was found that the 16 solo activities to prepare for TOEFL iBT speaking test were more frequently practised than the interactive speaking activity. “Talk to people in English” was the least frequently used ( $M=2.54$ ,  $SD=1.293$ ), and it was also considered the second least useful ( $M=3.49$ ,  $SD=1.016$ ). “Take notes while listening or reading” was not only the most frequently practised but also considered the most useful. Overall, the majority of the means of perceived usefulness were close to 4.0 (i.e., very useful). In terms of frequency of use, all of these activities were close to 3.0 (i.e., weekly), with three activities (“Take notes while listening or reading”, “Read aloud”, and “Increase listening input”) even close to 4.0 (i.e., once a day). Not surprisingly, due to the nature of intensive test preparation since the majority of the participants would be taking the test within six months (see *Time management and commitment*), all the 17 test preparation activities were frequently practised, though with some noticeable variation – some were more popular than others.

**Table 4 Frequency and usefulness of seventeen test preparation activities**

Test preparation activities	Frequency of use			Usefulness		
	N	M	SD	N	M	SD
Take notes while listening or reading	1374	3.76	1.085	1317	3.95	.953
Read aloud (incl. read after recording)	1397	3.72	1.069	1331	3.81	.978
Increase listening input	1372	3.62	1.030	1306	3.84	.930
Talk to myself on a given topic	1386	3.28	1.252	1312	3.68	.980
Practise speaking logically by using outlines, examples/details	1369	3.25	1.207	1299	3.93	.979
Practise timed speaking as in test	1365	3.21	1.207	1294	3.86	.983
Enhance reading ability	1367	3.21	1.146	1307	3.67	.956
Learn about TOEFL iBT Speaking related topics	1366	3.19	1.099	1300	3.79	.935
Summarize orally	1379	3.17	1.202	1303	3.77	.944
Do TOEFL iBT mock Speaking tasks	1367	3.09	1.145	1301	3.86	.986
Memorize sentence templates	1368	2.91	1.271	1294	3.49	1.075
Practise speaking to microphone or computer	1362	2.80	1.304	1292	3.59	1.074

Study TOEFL iBT Speaking rating criteria	1365	2.72	1.125	1296	3.66	1.085
Study TOEFL iBT ji-jing (other people's test experience)	1375	2.69	1.297	1290	3.64	1.081
Record my speaking to self-assess	1374	2.66	1.310	1289	3.62	1.083
Memorize model essays	1364	2.61	1.266	1282	3.33	1.118
Talk to people in English	1381	2.54	1.293	1299	3.49	1.016
Valid N (Listwise)	1301			1201		

Note: maximum point = 5

It is worth noting, “memorize model essays”, “memorize sentence templates” and “study TOEFL iBT ji-jing” were not among the most frequently practised test preparation activities, and the two memorization activities were in fact among the three least useful (the least useful activity being “memorize model essays” (M=3.33, SD=1.118).

No statistically significant difference between male and female students was observed in frequency of use of any of the 17 preparation activities. However, there were statistically significant differences between 16-18 and 19-24 year olds in 5 activities (with small effect sizes, see Table 5). Although overall “talk to people in English” was the least practised, the younger students (16-18 year olds) were more likely to “talk to people in English” than the older students (19-24 year olds) who tended to do more solo activities such as: study TOEFL iBT ji-jing, record own speaking to self-assess, practise speaking to microphone or computer, and practise timed speaking as in test. The younger students seemed to be more “interactive” in their preparation for the Speaking tasks, while the older students tended to do more frequently the activities that they can be in the driving seat.

**Table 5 Significant difference in frequency of use of test preparation activities (16-18 vs. 19-24 year olds)**

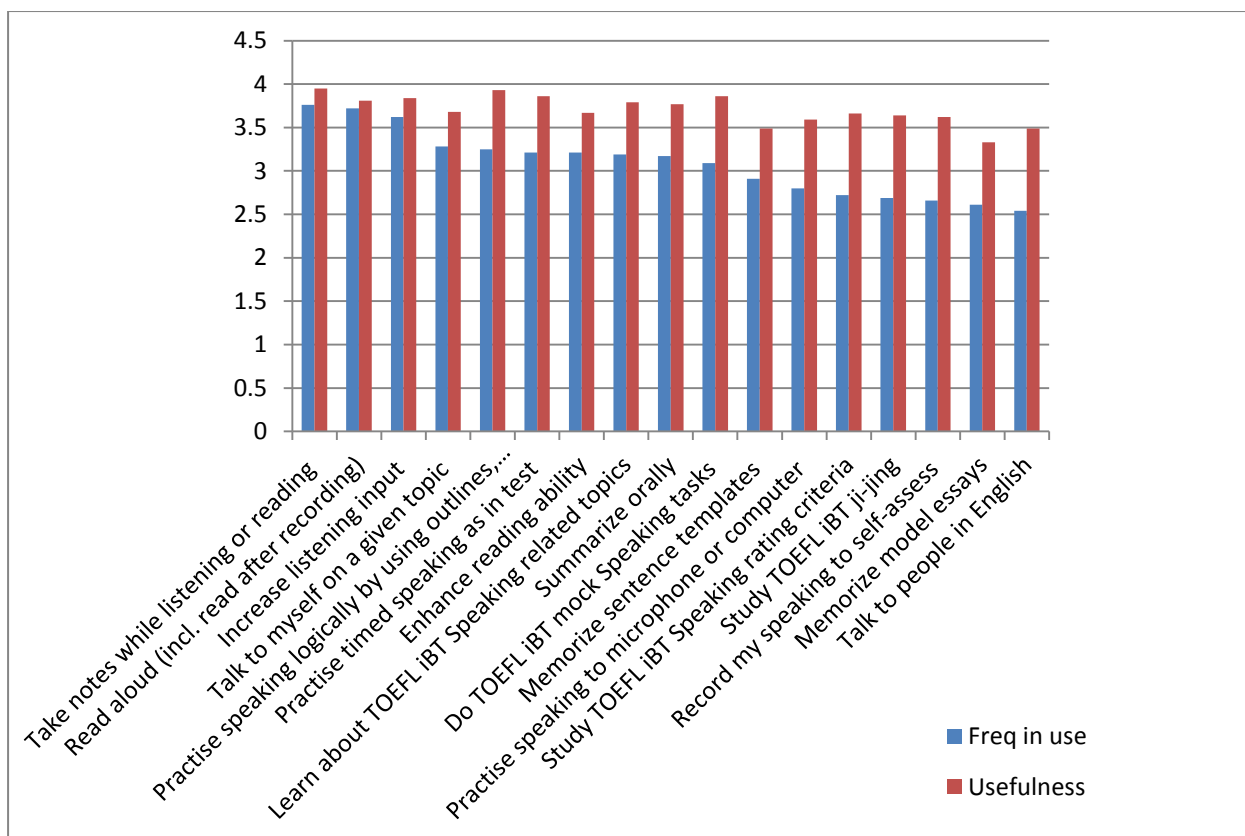
Frequency of use	t	df	Sig. (2-tailed)	Cohen's d	Mean: 16-18	Mean: 19-24	Mean Difference
Talk to people in English	5.286	1223	.000	.313	2.79	2.39	.399
Study TOEFL iBT ji-jing	-2.830	1216	.005	.170	2.56	2.78	-.217
Record my speaking to self-assess	-4.478	1217	.000	.268	2.48	2.83	-.346
Practise speaking to microphone or computer	-3.025	1209	.003	.178	2.67	2.90	-.233
Practise timed speaking as in test	-2.045	1208	.041	.117	3.14	3.28	-.146

Similarly, as shown in Table 6, those who would take TOEFL iBT to apply for undergraduate programmes practised “talk to people in English” and “memorize sentence templates” more frequently than those who would apply for Master programmes. The students applying for Master programmes however did more in recording their speaking to self-assess and practising speaking to microphone or computer. Among these differences, “talk to people in English” has the largest effect size (though still small according to Cohen’s d).

**Table 6 Significant difference in frequency of use of test preparation activities (applying for undergraduate vs. Master programmes)**

Frequency of use	t	df	Sig. (2-tailed)	Cohen's d	Mean: UG	Mean: Master	Mean Difference
Talk to people in English	5.601	1066	.000	.346	2.81	2.37	.437
Memorize sentence templates	2.019	1060	.044	.119	2.99	2.84	.156
Record my speaking to self-assess	-4.122	1062	.000	.246	2.51	2.83	-.329
Practise speaking to microphone or computer	-3.013	1054	.003	.185	2.68	2.92	-.241

Overall, the perceived usefulness of all test preparation activities was statistically significantly higher than the frequency of use (see Figure 1 below, and Table 4; t-test statistics available from the first author upon request).



**Figure 1 Frequency and usefulness of test preparation activities**

Note: maximum point = 5

Compared to the frequency of use, we found a larger number of test preparation activities had significant differences in their perceived usefulness between genders, age groups, and test-taking purposes. In six of the 17 activities, female students considered them significantly more useful than male students (see Table 7). There were eight activities that were significantly different between 16-18 and 19-24 year olds (see Table 8), and almost the same eight activities between test-taking purposes (see Table 9).

**Table 7 Significant difference in usefulness of test preparation activities (female vs. male)**

Usefulness	t	df	Sig. (2-tailed)	Cohen's d	Mean: female	Mean: male	Mean Difference
Talk to myself on a given topic	2.201	1262	.028	.123	3.74	3.62	.121
Study TOEFL iBT Speaking rating criteria	2.240	1249	.025	.121	3.73	3.60	.136
Record my speaking to self-assess	1.978	1240	.048	.120	3.69	3.56	.121
Practise timed speaking as in a test	2.647	1247	.008	.154	3.95	3.80	.146

Learn about TOEFL iBT Speaking related topics	2.387	1252	.017	.128	3.85	3.73	.126
Practise speaking logically by using outlines, examples/ details	2.432	1251	.015	.133	3.99	3.86	.134

Note: None of these activities was significantly different **between** female and male students in terms of their frequency of use.

**Table 8 Significant difference in usefulness of test preparation activities (16-18 vs. 19-24 year olds)**

Usefulness	t	df	Sig. (2-tailed)	Cohen's d	Mean: 16-18	Mean: 19-24	Mean Difference
Read aloud (incl. read after recording)	-2.779	1180	.006	.174	3.72	3.89	-.164
Talk to myself on a given topic	-2.731	1163	.006	.163	3.57	3.73	-.163
Talk to people in English	2.654	1149	.008	.159	3.59	3.43	.163
Study TOEFL iBT ji-jing (other people's test experience)	-2.062	1142	.039	.131	3.58	3.72	-.135
Study TOEFL iBT Speaking rating criteria	-2.011	1148	.045	.129	3.58	3.72	-.133
Record my speaking to self-assess	-4.959	1143	.000	.309	3.43	3.76	-.325
Practise speaking to my microphone or computer	-4.072	1145	.000	.243	3.43	3.69	-.267
Practise timed speaking as in a test	-4.139	1147	.000	.248	3.73	3.97	-.245

Note: The highlighted activities were also significantly different in their frequency of use between 16-18 and 19-24 year olds, see Table 5.

**Table 9 Significant difference in usefulness of test preparation activities (applying for undergraduate vs. Master programmes)**

Usefulness	t	df	Sig. (2-tailed)	Cohen's d	Mean: UG	Mean: Master	Mean Difference
Read aloud (incl. read after recording)	-2.704	1027	.007	.166	3.75	3.91	-.163
Talk to myself on a given topic	-2.320	1007	.021	.155	3.60	3.75	-.142
Talk to people in English	4.033	1001	.000	.261	3.62	3.36	.255
Summarize orally	-2.853	1007	.004	.182	3.66	3.83	-.168
Study TOEFL iBT ji-jing (other people's test experience)	-1.961	993	.050	.130	3.57	3.71	-.134
Record my speaking to self-assess	-5.252	991	.000	.337	3.42	3.78	-.357
Practise speaking to microphone or computer	-3.796	997	.000	.243	3.43	3.69	-.258
Practise timed speaking as in a test	-3.426	996	.001	.217	3.75	3.96	-.210

Note: The highlighted activities were also significantly different in their frequency of use between undergraduate and Master programmes, see Table 6.

### Course-books, websites and other materials

Learning materials play an integral role in test preparation; they define to a large extent what students can do in test preparation. In this section, we report what materials were available, how often they were used and how useful they were from the perspectives of both students and teachers.

*Official Guide (OG)*, *TOEFL Practice Online (TPO)* and *Delta's Key to the Next Generation TOEFL Test* were the three most frequently used materials ( $M \approx 3.0$ , i.e., weekly), and they were also considered the most useful by both students (see Table 10) and teachers (see Table 16). On average, *TPO* had the highest rating of usefulness ( $M=4.02$ ,  $SD=1.068$ , very useful).

**Table 10 Test preparation course-books (student data)**

	Frequency of use			Usefulness		
	N	M	SD	N	M	SD
Official Guide (OG)	1330	3.29	1.273	1222	3.89	1.075
TOEFL Practice Online (TPO)	1301	2.89	1.367	1168	4.02	1.068
Delta's Key to the Next Generation TOEFL Test	1282	2.89	1.468	1120	3.42	1.157
Barron's How to Prepare	1255	2.04	1.345	1039	3.01	1.209
Longman Preparation Course for TOEFL iBT	1249	1.99	1.346	1021	2.86	1.183
TOEFL iBT test sampler	1241	1.93	1.269	1005	3.04	1.213
TOEFL Value packs	1241	1.79	1.223	994	2.90	1.202
Kaplan TOEFL iBT	1235	1.57	1.104	976	2.61	1.175
Cambridge Preparation for the TOEFL Test	1233	1.49	1.054	970	2.57	1.196
Thomson The Complete Guide	1232	1.44	.994	965	2.53	1.178
Valid N (listwise)	1214			925		

#### Notes:

- The reduced number of cases/students in "usefulness" was due to the fact that some participants who chose the lowest frequency of use (i.e., never used) did not respond to the related question on "usefulness".
- Note: maximum point = 5

There were a number of significant differences in frequency of course-book use and their perceived usefulness between genders, age, and test-taking purposes. In frequency of use, the only significant difference between female and male students was in *Delta's Key to the Next Generation TOEFL Test* ( $M_{\text{female}}=2.97$ ,  $SD=1.454$ ,  $n=626$ ;  $M_{\text{male}}=2.79$ ,  $SD=1.478$ ,  $n=610$ ;  $t=2.152$ ,  $df=1234$ ,  $p<.0325$ ; **Cohen's  $d=.123$** ). However, more course-books had significant differences in frequency of use between 16-18 and 19-24 year olds (see Table 11), and between those applying for undergraduate and Master programmes (see Table 12). The older students

used the top three materials (*OG*, *TPO* and *Delta*) more frequently than the younger students.

**Table 11 Significant difference in frequency of use of course-books (16-18 vs. 19-24 year olds)**

Frequency of use	t	df	Sig. (2-tailed)	Cohen's d	Mean: 16-18	Mean: 19-24	Mean Difference
Official Guide (OG)	-4.164	1177	.000	.252	3.11	3.43	-.320
TOEFL Practice Online (TPO)	-3.427	1151	.001	.213	2.71	3.00	-.286
Delta's Key to the Next Generation TOEFL Test	-4.154	1135	.000	.252	2.68	3.05	-.378
Kaplan TOEFL iBT	2.584	1092	.010	.157	1.66	1.49	.176
Cambridge Preparation for the TOEFL Test	2.138	1090	.033	.134	1.56	1.42	.140
Thomson The Complete Guide	2.558	1089	.011	.153	1.52	1.37	.158

**Table 12 Significant difference in frequency of use of course-books (applying for undergraduate vs. Master programmes)**

Frequency of use	t	df	Sig. (2-tailed)	Cohen's d	Mean: UG	Mean: Master	Mean Difference
Official Guide (OG)	-3.932	1032	.000	.239	3.13	3.43	-.308
TOEFL Practice Online (TPO)	-3.943	1014	.000	.248	2.72	3.06	-.341
Longman Preparation Course for TOEFL iBT	2.750	972	.006	.179	2.11	1.87	.236
Delta's Key to the Next Generation TOEFL Test	-4.101	991	.000	.261	2.68	3.06	-.380
Kaplan TOEFL iBT	3.981	960	.000	.250	1.74	1.46	.288
Cambridge Preparation for the TOEFL Test	3.841	958	.000	.242	1.65	1.39	.267
Thomson The Complete Guide	4.651	956	.000	.307	1.63	1.32	.304

In terms of perceived usefulness of the test preparation course-books, *TPO*, *Delta* and *Barron* were rated higher by female than male students (see Table 13). *OG*, *TPO* and *Delta* were rated higher by older (19-24 year olds) than younger students (see Table 14 and Table 15), while *Kaplan*, *Cambridge* and *Thomson* were rated higher by those applying for undergraduate programmes (see Table 15).

**Table 13 Significant difference in usefulness of course books (female vs. male)**

Usefulness	t	df	Sig. (2-tailed)	Cohen's d	Mean: female	Mean: male	Mean Difference
TOEFL Practice Online (TPO)	2.286	1123	.022	.141	4.11	3.96	.145

Delta's Key to the Next Generation TOEFL Test	2.444	1078	.015	.148	3.50	3.33	.171
Barron's How to Prepare	2.380	1000	.018	.149	3.11	2.93	.181

**Table 14 Significant difference in usefulness of course books (16-18 vs. 19-24 year olds)**

Usefulness	t	df	Sig. (2-tailed)	Cohen's d	Mean: 16-18	Mean: 19-24	Mean Difference
Official Guide (OG)	-3.175	1080	.002	.197	3.78	3.99	-.216
TOEFL Practice Online (TPO)	-4.114	1031	.000	.267	3.88	4.16	-.282
Delta's Key to the Next Generation TOEFL Test	-2.809	990	.005	.189	3.31	3.53	-.218

**Table 15 Significant difference in usefulness of course books (applying for undergraduate vs. Master programmes)**

Usefulness	t	df	Sig. (2-tailed)	Cohen's d	Mean: UG	Mean: Master	Mean Difference
Official Guide (OG)	-2.188	943	.029	.140	3.80	3.95	-.154
TOEFL Practice Online (TPO)	-5.756	903	.000	.387	3.83	4.23	-.398
Delta's Key to the Next Generation TOEFL Test	-2.107	860	.035	.146	3.33	3.50	-.168
Kaplan TOEFL iBT	2.345	753	.019	.178	2.74	2.53	.202
Cambridge Preparation for the TOEFL Test	3.293	749	.001	.243	2.74	2.45	.287
Thomson The Complete Guide	2.481	746	.013	.179	2.66	2.45	.214

The teacher questionnaire data confirmed the popularity and usefulness of the top three identified in the student questionnaire data, i.e., *OG*, *TPO* and *Delta*. As test preparation is such a big and lucrative market, many teacher-turned authors have also written course-books, often using the brand name of big test preparation companies that they are/were affiliated with. In the teacher questionnaire, we asked teachers how often they used some of these Chinese books which have TOEFL iBT Speaking in their titles or as their main focus, and how useful they were. Two Chinese books (新托福口语金牌教程, 新托福口语黄金八十题) were found quite popular and useful.



**Table 16 Test preparation course-books (teacher data)**

	Frequency of use			Usefulness		
	N	M	SD	N	M	SD
Official Guide (OG)	22	3.86	1.125	16	4.06	1.181
TOEFL Practice Online (TPO)	22	3.82	1.368	16	4.25	1.125
TOEFL Value Pack plus	22	2.14	1.167	15	2.60	1.242
TOEFL Value Pack prep	21	2.19	1.327	14	2.57	1.284
ETS Pronunciation in English	22	2.05	1.133	15	2.33	.816
TOEFL iBT sample questions	22	3.05	1.397	15	2.80	.941
TOEFL iBT test sampler	22	2.91	1.377	15	2.87	1.060
TOEFL iBT Test Tips	22	2.41	1.221	15	2.53	.915
Longman Preparation Course for TOEFL iBT	21	2.52	1.327	14	2.50	.855
Delta's Key to the Next Generation TOEFL Test	21	3.14	1.424	16	2.94	1.237
Barron's How to Prepare	21	2.90	1.091	16	2.94	1.124
Kaplan TOEFL iBT	21	2.48	1.289	15	2.87	1.125
Cambridge Preparation for the TOEFL Test	21	2.00	1.049	15	2.80	1.082
Thomson The Complete Guide	20	2.00	1.170	14	2.71	1.069
Princeton Review's Cracking the TOEFL	21	1.90	1.136	15	2.60	.986
NOVA's Speaking and Writing Strategies for TOEFL iBT	21	1.67	.730	15	2.27	.884
新托福考试完全攻略(新东方, 张洪伟)	21	1.95	1.117	15	2.27	.704
新托福考试速战速决(新航道)	21	1.95	.921	15	2.27	.884
新托福考试口语特训(新东方, 李志研)	21	2.19	1.123	15	2.33	.900
TOEFL iBT 口语满分模板(新东方, 邱政政)	21	2.00	1.049	15	2.20	.775
新托福考试口语胜经(新东方, 翟少成)	21	2.24	1.411	15	2.33	1.047
新托福口语金牌教程(新航道)	21	2.95	1.596	15	2.60	1.298
新托福口语黄金八十题	21	2.67	1.390	15	2.73	1.163
Valid N (listwise)	20			13		

Note: maximum point = 5

The interviews and classroom observation data also showed additional materials that were used, including several “vocabulary books” (mainly Chinese publications), videos from Friends, VOA, BBC Documentary, Scientific America, CNN and TED, and computer software for recording and timing (e.g., Cool Edit, Adobe Audition). Furthermore, we found that nearly all of the learning materials were available to download/access free of charge. Apparently this is a huge resource for test takers and it is particularly welcome by teachers and students; however there does seem to be some perennial copyright infringement. It is interesting to note that many of the international course-books (e.g., *Delta*, *Kaplan*, *Longman*) are also published specifically for the Chinese market in collaboration with big test preparation companies (e.g., New Oriental, New Channel).

In addition to the published course-books, students and teachers mentioned at the interviews that some popular websites also provide excellent resources for test news, test preparation materials (including ji-jing) and other opportunities for them to improve their Speaking ability. These websites included those we already reported in *Instruments and procedure*, plus the following:

- [www.chasedream.com](http://www.chasedream.com)
- [www.putclub.com](http://www.putclub.com)
- [www.hjenglish.com](http://www.hjenglish.com)
- [ke.qq.com](http://ke.qq.com)
- [www.51test.net](http://www.51test.net)
- [www.ets.org/toefl](http://www.ets.org/toefl)
- [www.nytimes.com](http://www.nytimes.com)
- [www.bbc.co.uk/news/world](http://www.bbc.co.uk/news/world)

Increasingly now there are Apps available for TOEFL iBT test preparation (e.g., Xiaoma TOEFL, <https://itunes.apple.com/us/app/xiao-ma-tuo-fu/id790626096?mt=8>).

### Reasons and expectations for taking test preparation courses

The vast majority of the participants were attending intensive test preparation courses. As shown in Table 17, there appeared to be three distinct reasons and expectations for taking intensive test preparation courses. We call the first factor: *to learn through speaking activities and tasks* within the classroom context so that language proficiency, academic study skills and test-taking confidence can be boosted; the second factor: *to learn about the test features and test-taking strategies* so that language proficiency, academic skills and test-taking confidence can be boosted; and the third factor: *to enhance the social aspects of test-taking*, e.g., self-confidence, friendships and parents' expectations. Improving confidence and reducing test fear were common across the three factors; and improving language proficiency and academic study skills were shared by the first and the second factor. It seems that doing speaking activities/tasks and learning about the test features and test-taking strategies were both considered capable of playing an integral role in improving confidence and therefore reducing test fear. The underlying construct of the reasons and expectations of female students taking test preparation courses was very similar to that of male students (statistics available from the first author upon request).

**Table 17 Rotated component matrix: reasons and expectations for taking test preparation courses**

Reasons and expectations for taking test preparation courses	Component		
	1	2	3
to learn test taking strategies		.856	
to learn test formats		.881	
to learn test topics		.853	
to improve English language proficiency	.440	.460	
to improve academic study skills	.372	.469	

hoping teachers will correct my pronunciation/intonation	.814		
hoping teachers will correct my grammatical mistakes	.805		
hoping teachers will organise speaking practice tasks	.710		
hoping teachers will predict speaking tasks	.539		
hoping to have more opportunities to speak English at class	.835		
hoping to learn from classmates and improve together	.763		
to improve confidence	.416	.445	.369
to reduce test fear	.355	.390	.486
to make new friends			.803
to meet parents' expectations			.840
Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .868			
Bartlett's Test of Sphericity Approx. Chi-Square = 8893.810, df=105, sig.<.0005			
Extraction Method: Principal Component Analysis			
Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 4 iterations.			
Three factors were extracted, based on Eigenvalues greater than 1.			
Absolute value of coefficient below 0.30 was suppressed in this table.			
Analysis N=1098; Rotation sums of squared loadings: 62.59%			

As shown in Table 18, the top five reasons and expectations for taking intensive preparation courses were: to learn test-taking strategies (M=4.40, SD=.864), to learn test formats (M=4.36, SD=.922), to improve confidence (M=4.23, SD=1.031), to learn test topics (M=4.21, SD=.992), and to practise speaking tasks organized by teachers (M=4.21, SD=1.00). Within the top five reasons, the students considered it a lot more important to learn test-taking strategies and test formats than to practise speaking tasks in lessons (test-taking strategies vs. speaking activities,  $t=5.413$ ,  $df=1151$ ,  $p<.0005$ ; test formats vs. speaking activities,  $t=4.006$ ,  $df=1148$ ,  $p<.0005$ ). It is also interesting to observe that the top five reasons had smallest standard deviations, which indicates a stronger agreement among the students in their views on the top five reasons than in the other 10 reasons which had larger standard deviations and smaller means. **There was no significant difference in any of the top five reasons between genders, age (except for “to learn test topics”,  $t=-2.184$ ,  $df=1066$ ,  $p<.0295$ ), and test-taking purposes.**

**Table 18 Reasons and expectations for taking test preparation courses (student data)**

Reasons and expectations for taking test preparation courses	N	Mean	Std. Deviation
to learn test taking strategies	1186	4.40	.864
to learn test formats	1182	4.36	.922
to learn test topics	1182	4.21	.992
to improve English language proficiency	1187	4.17	1.081
to improve academic study skills	1173	3.95	1.172

teachers will correct my pronunciation/intonation	1220	3.71	1.339
teachers will correct my grammatical mistakes	1214	3.69	1.298
teachers will organise speaking practice tasks	1219	4.21	1.000
teachers will predict speaking tasks	1211	3.70	1.283
to have more opportunities to speak English at class	1210	3.67	1.280
to learn from classmates and improve together	1214	3.79	1.215
to improve confidence	1222	4.23	1.031
to reduce test fear	1204	4.06	1.149
to make new friends	1195	3.22	1.433
to meet parents' expectations	1192	2.85	1.542

Note: maximum point = 5

We also asked the teachers why they thought their students were attending test preparation courses. It was found the teachers' views were broadly in line with the students' (see Table 19). The top five reasons and expectations that the teachers identified were similar to students': (i) to learn test formats, and to reduce test fear (M=4.64 for both), (iii) to learn speaking test tasks that teachers will predict (M=4.62), (iv) to learn test taking strategies (M=4.50), (v) to learn test topics, to practise speaking tasks that teachers organize, and to improve confidence (M=4.45 for all the three reasons/expectations). However, there was some slight difference in the order of importance, and one notable difference between teachers and students. That is, teachers thought it was one of the student's top priorities to learn about the speaking test tasks they would predict; however, students did not seem to anticipate that as much as teachers do (c.f. Table 18).

**Table 19 Reasons and expectations for taking test preparation courses (teacher data)**

Reasons and expectations for taking test preparation courses	N	Mean	Std. Deviation
to learn test taking strategies	22	4.50	1.058
to learn test formats	22	4.64	.727
to learn test topics	22	4.45	.858
to improve English language proficiency	22	4.23	.922
to improve academic study skills	22	3.68	1.211
teachers will correct my pronunciation/intonation	22	4.23	1.020
teachers will correct my grammatical mistakes	22	4.18	1.053
teachers will organise speaking practice tasks	22	4.45	.912
teachers will predict speaking tasks	21	4.62	.865
to have more opportunities to speak English at class	22	4.32	1.041
to learn from classmates and improve together	22	4.00	.926

to improve confidence	22	4.45	.912
to reduce test fear	22	4.64	.790
to make new friends	22	3.50	1.058
to meet parents' expectations	22	3.68	1.041

Note: maximum point = 5

### Features of test preparation courses

In this section, we report the features of the intensive test preparation courses, in terms of the average number of students in classroom, characteristics of the teaching force (including teacher qualifications, language proficiency, experience, and knowledge about TOEFL iBT), medium of instruction, and the test preparation activities that teachers organise during TOEFL iBT Speaking lessons.

Although smaller class (including one-to-one tuition) has become increasingly popular now in China, large class still seemed to be most common, nearly 44% of the participants said they attended courses which had more than 100 students (see Table 20).

**Table 20 Number of students in a Speaking lesson**

Number of students in a Speaking lesson	Frequency	Valid %
1 student (myself)	10	.8
2-5 students	32	2.5
6-10 students	75	5.9
11-20 students	166	13.1
21-40 students	221	17.5
41-100 students	205	16.2
more than 100 students	554	43.9
Valid N	1263	

The teachers were young, energetic, well-educated, very confident, highly proficient in English, and mainly on short-term contracts. Of the 23 teacher participants, around 40% was below 25 and 40% between 26-30 years old; 17 are female and 6 male; 57% had a Bachelor and 40% a Master degree, and one teacher had a doctoral degree; 70% of them earned their degree(s) from Chinese universities, and 26% from overseas. Their employment history with their current test preparation school varied greatly, from 1 month to 24 years; 22% (5 of them, which was the mode) had worked with their current employer for about 12 months. The majority of them (74%) were in full-time employment and 26% part-time.

Their experience in teaching English ranged from 6 months to 24 years ( $M=66.39$  months,  $SD=75.59$ ); and their experience in teaching TOEFL ranged from 2 months to 10 years ( $M=28.74$  months,  $SD=28.38$ ); three of them had experience in teaching the old TOEFL. More than half of them (61%) received some training on how to teach TOEFL iBT (either internally or via ETS training courses); and the majority (70%) also taught other TOEFL iBT courses (e.g., Writing, Reading, Listening, Grammar and Vocabulary). This was pretty consistent with the student data. According to the student data, the majority of their Speaking teachers (69.4%) also taught other skills/courses that they attended; only 30.6% of the students reported that their teachers taught them TOEFL iBT Speaking course only.

The teachers reported that they were required by their employer, as part of their job, to take TOEFL iBT in order to get the first-hand experience of the test (see Table 21); 35% had taken the test more than once. The teacher participants were highly proficiency in English – the mean of their TOEFL iBT Speaking score was 27.61 ( $SD=1.75$ ,  $min=24$ ,  $max=30$ ); they were generally very confident about their knowledge about TOEFL iBT Speaking tasks and how to teach and assess them (see Table 22).

**Table 21 Purpose of taking TOEFL iBT test (teacher data)**

Purpose of taking TOEFL iBT test	N	Mean	Std. Deviation
to learn test format	19	4.84	.688
to learn potential test topics	19	4.68	.671
to learn test difficulty level	19	4.79	.631
to learn time management during test	19	4.79	.419

Note: maximum point = 5

**Table 22 Teachers' self-assessment of their knowledge about TOEFL iBT Speaking**

Teachers' self-assessment of	N	Mean	Std. Deviation
knowledge of different methods for teaching speaking	21	4.52	.602
knowledge of different methods for assessing speaking	21	4.14	.964
knowledge of TOEFL iBT independent Speaking tasks	21	4.62	.740
knowledge of TOEFL iBT integrated Speaking tasks	21	4.43	.870
knowledge of TOEFL iBT rating criteria for independent Speaking tasks	21	4.48	.814
knowledge of TOEFL iBT rating criteria for integrated Speaking tasks	21	4.43	.746

Note: maximum point = 5

We asked the teachers (n=14) who also had experience in teaching for other international speaking tests, and those who (n=13) also had experience in teaching general speaking courses (i.e., not related to any high-stakes test) to compare the demand on teachers to teach these courses. As shown in Table 23, the majority of them said that it was a lot more demanding to teach TOEFL iBT speaking, in several aspects, e.g., time for lesson planning, teaching resources, teachers' own language proficiency and teaching skills. In terms of the overall demand, 78.6% of these teachers thought it more demanding to teach TOEFL iBT speaking than other international speaking tests, and 84.6% of them thought it more demanding than to teach general speaking courses.

**Table 23 Comparison of demand of teaching TOEFL iBT with other speaking courses**

	Higher than (%)		Lower than (%)		Similar to (%)	
	Other international tests	General speaking course	Other international tests	General speaking course	Other international tests	General speaking course
Overall demand on teacher	78.6	84.6	14.3	0	7.1	15.4
Time for lesson planning	85.7	76.9	7.1	7.7	7.1	15.4
Teaching resources	78.6	76.9	14.3	7.7	7.1	15.4
Teachers' language proficiency	71.4	76.9	14.3	0	14.3	23.1
Teaching methods and skills	64.3	84.6	7.1	7.7	28.6	7.7

Below we report what happened in the intensive TOEFL iBT speaking courses, from both teachers and students perspectives. In terms of medium of instruction, 14.3% of the teachers said they used “mainly English”, and 85.7% “half English, half Chinese” (valid N=21). According to the student data, 24.2% of the students said their speaking lessons were mainly in English, 52.9% of them said their lessons were “half English, half Chinese”, and 21.4% of them said their lessons were mainly in Chinese (valid N=1262). According to the lessons we observed, it is fair to say that the majority of the speaking lessons were largely “half English, half Chinese”.

According to the student data, the classroom activities that their teachers most frequently organised were more related to explaining test-taking strategies than actually facilitating students to perform speaking tasks (see Table 24). In the students' view, the top five most frequently used

and also most useful classroom activities were: teachers explaining test-taking strategies for independent and integrated speaking tasks, explaining how to improve overall speaking proficiency, assigning homework, and doing mock tests. The more frequently a classroom activity was organised by teachers, the more useful the students thought it was (Note: this strong correspondence between frequency of use and perceived usefulness was also observed in other test preparation activities and course books).

**Table 24 Classroom activities organised by teachers (student data)**

Classroom activities organised by teachers	Frequency of use			Usefulness		
	N	M	SD	N	M	SD
explain how to improve overall speaking proficiency	1178	4.04	.850	1113	4.08	.847
explain strategies for independent tasks	1178	4.19	.766	1114	4.16	.816
explain strategies for integrated tasks	1177	4.16	.778	1110	4.16	.814
provide sample/models for independent tasks	1169	3.74	1.021	1106	3.87	.968
provide sample/models for integrated tasks	1169	3.70	1.026	1102	3.84	.980
explain/study rating scales (overall)	1172	3.89	.942	1108	3.99	.929
do mock tests	1170	3.97	.959	1106	4.08	.894
evaluate student performance (e.g., in content & structure)	1173	3.75	1.043	1106	3.89	.938
correct pronunciation	1176	3.54	1.121	1101	3.81	.962
correct grammatical errors	1170	3.47	1.141	1100	3.73	1.004
study ji-jing (other people's test experience)	1163	3.00	1.322	1079	3.65	1.087
assign homework	1168	3.99	.945	1094	4.04	.913
organise speaking activities other than mock tests	1165	3.21	1.291	1093	3.55	1.093
Valid N (listwise)	1127			1036		

Note: maximum point = 5

We examined the teacher data (see Table 25) to see if there was any difference between teachers and students in their views about the intensive test preparation courses. The teacher data showed that the top three classroom activities that teachers did were: encouraging students to speak, organising speaking activities other than mock tests, and evaluating students' speaking task performance. Their next five activities were exactly the same as the students' top five (see Table 24).

**Table 25 Classroom activities organised by teachers (teacher data, part 1)**

Classroom activities organised by teachers	Frequency of use			Usefulness		
	N	M	SD	N	M	SD
explain how to improve overall speaking proficiency	22	4.14	.710	20	4.20	.834
explain strategies for independent tasks	22	4.36	.658	20	4.45	.686



explain strategies for integrated tasks	22	4.18	.907	20	4.35	.671
provide sample/models for independent tasks	22	3.73	1.120	20	3.70	.865
provide sample/models for integrated tasks	20	3.95	.999	19	3.74	.872
explain rating scales (independent tasks)	20	4.00	.918	18	3.89	.758
explain rating scales (integrated tasks)	21	4.05	.921	19	4.00	.745
explain/study rating scales (overall)	21	3.95	.865	19	3.84	.898
do mock tests	22	4.27	.767	19	4.37	.761
evaluate student performance (e.g. in content & structure)	21	4.43	.676	19	4.21	.713
correct pronunciation	22	3.91	.750	19	3.84	.834
correct grammatical errors	21	3.81	.873	19	3.84	.834
study ji-jing (other people's test experience)	21	3.57	1.028	18	3.78	.878
assign homework	21	4.38	.973	19	4.47	.841
organise speaking activities other than mock tests	21	4.52	.750	19	4.47	.772
encourage students to speak in class	21	4.71	.463	19	4.58	.607
Valid N (listwise)	18			16		

Note: maximum point = 5

We also asked the teachers to rate a variety of TOEFL iBT speaking or speaking-related activities: how often they used them in classroom and how useful they were for preparing for TOEFL iBT speaking test (see Table 26). The most frequently organised speaking activities included: practise timed speaking as in test (M=4.41), organising ideas by introduction, body and conclusion (M=4.36), making a point with supporting details and examples (M=4.32), connecting ideas from notes (M=4.27), writing an outline before talking (M=4.24), taking notes while listening (M=4.24), learning about TOEFL iBT Speaking related topics (M=4.18), summarising orally from listening (M=4.10), brainstorming for keywords before talking (M=4.05), and summarising orally from reading (M=4.00). These top 10 activities all had a mean  $\geq 4.00$  (the maximum possible score being 5.00), which indicates a high frequency of use. The teachers also considered these activities the top 10 most useful, in almost the same order as frequency of use, with the most useful activity being time management (i.e., practise timed speaking as in test). The 10 activities (except, summarising from listening, and summarising from reading) all aimed explicitly to help the students cope with the procedural and content requirement and demand of the TOEFL iBT Speaking test. As shown in Table 26, language learning activities (e.g., role play, group discussion, debate, read aloud, retell a story, listen to English movie, radio and TV programmes, read newspapers or magazines, summarise orally from reading and listening) that could enhance more directly the students' language proficiency than the instruction of test-taking strategies could, however, were less practised in classroom.

**Table 26 Classroom activities organised by teachers (teacher data, part 2)**

Classroom activities organised by teachers	Frequency of use			Usefulness		
	N	M	SD	N	M	SD
role play	22	2.59	.908	18	2.83	.857
one-to-one talk	22	3.68	.839	18	3.72	.895
group discussion	22	3.55	1.011	18	3.61	.916
Debate	22	3.00	1.345	18	3.44	.984
oral presentation/speech	22	3.64	1.093	18	3.56	.984
read aloud	22	2.68	1.171	17	2.71	1.105
read after recording	22	2.86	1.125	19	2.95	1.026
talk to myself on a given topic (with notes or outlines)	22	3.32	1.129	19	3.58	.961
talk to myself on a given topic (without notes or outlines)	21	2.95	.921	19	3.32	.749
retell a story	22	3.41	1.054	19	3.42	1.071
paraphrase a sentence	22	3.59	1.098	19	3.68	.820
summarize orally from reading	21	4.00	.775	18	3.94	.725
summarize orally from listening	21	4.10	.768	18	4.17	.618
practise speaking to microphone or computer	21	2.95	1.024	18	3.33	.840
record speaking to self-assess	22	3.32	1.129	19	3.79	.855
interpretation (oral translation)	22	3.05	1.214	19	3.26	1.147
take notes while listening	21	4.24	.831	18	4.28	.752
take notes while reading	21	3.95	1.024	18	3.94	.998
transcribe audio recordings word by word	21	2.81	1.123	18	3.11	1.023
memorize words relevant to Speaking tasks	22	3.82	1.097	19	4.05	.911
memorize sentence templates relevant to Speaking tasks	22	3.86	.990	19	3.89	1.049
memorize model essays	22	3.00	1.024	19	3.32	.946
learn about TOEFL iBT speaking related topics	22	4.18	.907	19	4.05	.911
practise timed speaking as in test	22	4.41	.796	19	4.42	.769
brainstorm for keywords before talking	22	4.05	1.046	19	4.05	1.079
write an outline before talking	21	4.24	.944	19	4.16	.958
practise connecting ideas from notes	22	4.27	.703	19	4.32	.820
practise making a point with supporting details/examples	22	4.32	.780	19	4.32	.820
practise organising ideas by introduction, body & conclusion	22	4.36	.790	19	4.32	1.057
listen to English radio programmes	21	3.05	1.117	18	3.44	1.149
watch English movie/TV programmes	22	3.41	.959	19	3.79	.855
read English newspaper/magazines	22	3.27	1.162	19	3.84	.898
Valid N (listwise)	19			15		

Note: maximum point = 5

### Test-taking strategies

As we reported in the previous three sections (*Frequency and usefulness of different test preparation activities*, *Reasons and expectations for taking test preparation courses*, *Features of test preparation courses*), learning test-taking strategies was among the most important reasons

that students had for attending intensive preparation courses, and teaching and learning test-taking strategies was indeed one of the core classroom activities **as we observed in the lessons**. From the teacher and student interviews and questionnaire data and recorded lessons, we identified two types of test-taking strategies or tricks that the students were being trained: procedural, and content-related. Below are some typical test-taking tips/strategies that teachers talk about in lessons and interviews.

*Procedural:*

- Good time management.
- Control speed and rhythm, pause where necessary.
- Slow down when necessary.
- Don't leave a big gap between sentences.
- Speak loud and fluently, stay calm and relaxed and be brave.
- Don't be shy, just try.
- Keep talking.
- Use shorthand to take notes.
- Make an outline before talking.
- During your break time, listen to what test takers next to you are talking about.

*Content-related:*

- Never speak in Chinese even if you are stuck.
- Make full use of what you've recited.
- Imitate American accent.
- Do not use too complex sentences; make it colloquial, simple but elegant.
- Use connectives to make logic arguments.
- Use more examples.
- Paraphrase what you've heard or read.
- If you can't understand the word, imitate the pronunciation.
- If you can't understand the listening input, try to listen as much as you can, and say only what you've heard.
- Try to say as much as what you've heard.
- Speak to fill the time, even it is nonsense.
- Say some "nonsense" sentences so you can find some time to plan.
- Do not use too many speech fillers such as uh, like and er.
- Do not use too many "I think..."

### Five-component test preparation: A holistic approach

In the previous sections, we examined separately the different test preparation activities, course books, reasons and expectations for attending intensive presentation courses, and features of test preparation courses. It is important that we take a holistic view to explore the underlying construct of test preparation. A principal component analysis was conducted on the student data, including all the 55 items/questions in relation to a range of test preparation activities outside and during lessons, course books, and reasons and expectations for taking intensive preparation courses.

**Table 27 Five components of test preparation**

	Component				
	1	2	3	4	5
Read aloud (incl. read after recording)	.552				
Talk to myself on a given topic	.582				
Talk to people in English	.389				
Take notes while listening or reading	.549				
Summarize orally	.660				
Memorize sentence templates	.545				
Memorize model essays	.500				
Do TOEFL iBT mock Speaking tasks	.725				
Study TOEFL iBT ji-jing	.594				
Study TOEFL iBT speaking rating criteria	.619				
Record my speaking to self-assess	.652				
Practise speaking to microphone or computer	.686				
Practise timed speaking as in test	.737				
Increase listening input	.541				
Enhance reading ability	.584				
Learn about TOEFL iBT speaking related topics	.685				
Practise speaking logically by using outlines, examples/details	.720				
Official Guide					.426
TOEFL Practice Online (TPO)	.380				
TOEFL Value packs				.760	
TOEFL iBT test sampler				.703	
Longman Preparation Course for TOEFL iBT				.632	
Delta's Key to the Next Generation TOEFL Test					.448
Barron's How to Prepare				.646	
Kaplan TOEFL iBT				.833	
Cambridge Preparation for the TOEFL Test				.868	
Thomson The Complete Guide				.873	
hoping to learn test taking strategies					.659
hoping to learn test formats					.651
hoping to learn test topics					.626
hoping to improve English language proficiency		.591			
hoping to improve academic study skills		.554			
hoping teachers will correct my pronunciation/intonation		.770			
hoping teachers will correct my grammatical mistakes		.775			
hoping teachers will organise a series of speaking practice tasks		.616			
hoping teachers will predict speaking tasks		.600			
hoping to have more opportunities to speak English at class		.776			
hoping to learn from classmates and improve together		.729			
helpful to improve my confidence		.597			
helpful to reduce my test fear		.586			
helpful to make new friends		.610			
helpful to meet parents' expectations		.473			

explaining how to improve overall speaking proficiency			.586		
explaining strategies for independent tasks			.683		.395
explaining strategies for integrated tasks			.674		.433
providing sample/models for independent tasks			.667		
providing sample/models for integrated tasks			.641		
explaining rating scales			.680		
doing mock tests			.618		
speaking activities other than mock tests			.504		
evaluating/commenting performance			.631		
correcting pronunciation			.643		
correcting grammatical errors			.640		
studying ji-jing			.514		
assigning homework			.609		

Kaiser-Meyer-Olkin Measure of sampling adequacy: .893  
Bartlett's test of sphericity approx. Chi-square: 26220.444, df=1485, sig.<.0005  
Extraction Method: Principal Component Analysis.  
Rotation Method: Varimax with Kaiser Normalization.  
Rotation converged in 6 iterations; Rotation sums of squared loadings (5 factors): 48.78%  
Analysis N=858. Absolute value of coefficients below 0.36 was suppressed.

**Table 28 Component transformation matrix**

Component	1	2	3	4	5
1	.605	.494	.507	.316	.182
2	.579	-.552	-.398	.442	-.086
3	-.369	.396	-.227	.692	-.419
4	.246	.531	-.728	-.268	.237
5	-.320	-.113	-.058	.392	.853

The principal component analysis indicated that a five-component structure can best explain for test preparation (see Table 27 and Table 28 above). **We call these five components:** (i) practising speaking or speaking-related activities outside classroom, (ii) anticipating academic as well as non-academic benefits (such as boosting confidence and reducing fear) from attending intensive preparation courses, (iii) doing or listening to classroom activities that teachers organize in preparation courses, (iv) studying course books and other learning materials, and (v) attempting to acquire test-taking strategies from Official Guide (OG) and other course books as well as teachers instructions during lessons.

### ***Relationship between test preparation and Speaking test performance***

We looked at two types of data to understand the relationship between test preparation and performance. The first source of data came from the participants' judgement. Both the teachers and the students believed that test preparation was very useful for improving test performance

( $M_{\text{teacher}}=4.18$ ,  $SD=.853$ ,  $N=22$ ;  $M_{\text{student}}=3.77$ ,  $SD=.779$ ,  $N=1200$ ). The second source of data came from the students' actual test performance. As reported in *Instruments and procedure*, we were only able to obtain test scores from a small number of participants (see Table 29).

**Table 29 TOEFL iBT test performance**

	N	Minimum	Maximum	Mean	Std. Deviation
Reading	207	7	30	23.47	5.069
Listening	205	2	30	20.80	5.563
Speaking	260	7	28	19.87	3.451
Writing	211	9	30	22.00	4.395
Total	293	33	112	84.57	14.945
Valid N (listwise)	205				

We examined the correlations between a student's TOEFL iBT Speaking test score and the frequency of him/her using each test preparation activity (see Table 4 and Table 24) and course-book (see Table 10). Eight preparation activities and one resource (TPO) were found to have statistically significant, but not so strong, correlations with Speaking test score (see Table 30). Only one preparation activity (do mock tests in classroom context) that teachers organised in classroom was significantly related to test performance.

**Table 30 Test preparation activities that have significant correlations with Speaking performance**

Significant individual variables	r.	ANOVA	
		F	sig.
Talk to myself on a given topic	.149	5.521	.020
Take notes while listening or reading	.135	4.487	.035
Summarize orally	.132	4.369	.038
Do TOEFL iBT mock Speaking tasks	.157	6.076	.014
Practise timed speaking as in a test	.173	7.402	.007
Learn about TOEFL iBT related topics	.123	3.715	.055 (borderline)
Practise speaking logically by using outlines, examples/ details	.224	12.729	.000
TOEFL Practice Online (TPO)	.232	13.249	.000
Do mock tests (in classroom context)	.141	4.180	.042

We then entered all these variables in the simple regression analysis; it was found that together they can account for 11.1% of TOEFL iBT Speaking test performance ( $F_{9, 182}=2.518$ ,  $p<.0105$ ). The frequency of using TOEFL Practice Online (TPO) was the only variable that remained capable of predicting the students' TOEFL iBT Speaking test score (see Table 31). TPO alone could explain about 5.4% of the variance in Speaking test score.

**Table 31 Simple linear regression**

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	16.741	1.333		12.560	.000
Talk to myself on a given topic	-.066	.219	-.024	-.303	.762
Take notes while listening or reading	.081	.273	.025	.296	.767
Summarize orally	-.009	.230	-.003	-.038	.970
Do TOEFL iBT mock Speaking tasks	.059	.307	.019	.194	.847
Practise timed speaking as in test	.086	.267	.029	.321	.749
Learn about TOEFL iBT speaking related topics	-.308	.292	-.090	-1.058	.292
Practise speaking logically by using outlines, examples/details	.417	.268	.143	1.554	.122
TOEFL Practice Online (TPO)	.626	.185	.260	3.378	.001
Do mock tests (in classroom context)	.096	.249	.030	.388	.699

## Discussion and conclusion

Preparation for high-stakes tests is prevalent in both formal and shadow education (i.e., educational activities outside of formal schooling), and it has widespread social and economic implications (Bray, 2007). The higher the stakes of the test, the stronger the urge to engage in specific test preparation practices that aim or claim to be able to enhance test performance. As Crocker eloquently synthesized, “No activity in educational assessment raises more instructional, ethical, and validity issues than preparation for large-scale, high-stakes tests” (Crocker, 2006, p.115), the contribution that research on “test preparation” can make towards understanding the instructional, ethical and validity issues of any educational assessment cannot be overstated. Developing the “validity argument” of any test is a holistic and ongoing process. Research on “test preparation” from test takers’ and teachers’ points of view should be built into such a process. We have argued for the importance of researching test preparation for test publishers and other stakeholders to understand the validity issues of the tests concerned, because not only test preparation itself but also any variation in intensity and types of test preparation could introduce construct-irrelevant variances in varying degrees (see Haladyna and Downing 2004). However, the nature of test preparation is far more complex than what many stakeholders, e.g., policy makers, test developers, teachers and students, have assumed. We have argued that test preparation must be investigated with a contextualized approach, taking into consideration the characteristics of test takers and teachers as the major players in test preparation, as well as the

local and global context where test preparation takes place.

Below we summarize and discuss the main findings of the project, aiming to demonstrate the multiple-components and dynamics of test preparation. What do test takers and teachers do when they prepare for TOEFL iBT speaking tasks? How useful are the different test preparation activities and materials from teachers and students perspectives? What are the relationships between test takers' TOEFL iBT test score and the preparation activities and materials they used? These are the three research questions that the data of this project can address (Note: RQ2 was dropped due to the insufficient number of test takers who were preparing for the test on their own).

***RQ 1 and RQ3: How do Chinese students and teachers prepare for TOEFL iBT speaking tasks, and how useful do they think test preparation is?***

To answer these two research questions, we asked a number of sub-questions with regard to students' test preparation experience to thematically present and discuss our main findings.

**When did students start to take intensive preparation courses, how did they manage their test preparation time and focus, and what did students do outside classroom during the intensive preparation period leading up to their test day?**

The majority of the students decided to take intensive preparation courses quite close to their planned or registered test date, hoping to achieve some quick and short-term fix/improvement in their test score. Only 10% of the participants prioritized speaking as No.1 in their preparation for TOEFL iBT. It was Listening that the majority of the students prioritized in their time allocation. Overall the most common ratio of time allocation to preparing for the speaking test was 20-40% of the total preparation, across the board (genders, age, and test-taking purposes). Test takers' specific focus on receptive skills or sections that require predominantly receptive skills (listening and reading) rather than productive skills (writing and speaking) indicates not only the language learning environment in China where the opportunities to speak and write in English for real communication purposes are still rare, but also the general belief among teachers and test takers that listening and reading test scores and abilities can be improved much more quickly than writing and speaking via intensive preparation. Focusing on listening and reading is a positive and prerequisite route towards improving performance in the TOEFL iBT speaking test which



does require both listening and reading skills in its integrated speaking tasks. It also indicates that the positive washback which Butler et al. (2000) anticipated the introduction of the speaking test would bring – test takers will “learn to communicate orally” – may take quite a while to be achieved. Test takers would naturally switch to what they could do within a particular learning context. In this case, the lack of opportunity to “talk to people in English” is evident.

Students did more solo speaking activities (i.e., what they could do on their own) than interactive speaking activity (e.g., talk to people in English). “Talk to people in English” was the least practised. Not only did the students practise most frequently note-taking during reading or listening – a key process of TOEFL iBT integrated speaking tasks – they also considered note-taking the most useful preparation for TOEFL iBT speaking test. “Read aloud” and “increase listening input” were the next two most frequently practised (nearly “once a day”), with the rest of preparation activities being practised close to “weekly”.

However, there are encouraging signs of change. The 16-18 year olds spent significantly more time preparing for Speaking skill than the 19-24 year olds, who allocated more time to prepare for Listening instead. The 16-18 year olds were also more likely than 19-24 year olds to “talk to people in English” when they prepared for their TOEFL iBT speaking test, while the 19-24 year olds did more test preparation activities that they can do on their own (i.e., without interaction with other people). These are encouraging signs of change of focus of China’s next generation of English language learners towards communicative language learning – the positive washback Butler et al. (2000) anticipated.

There are other encouraging signs that could be attributable to the introduction of the speaking test, especially the integrated speaking tasks which our participants found less coachable than the independent speaking tasks. Rote learning and memorization has traditionally been a label of Chinese learners; and topic-based monologue has long been considered highly coachable. We have heard that students were taught to memorize a short essay on “playing football” to answer a number of topic-based questions, for example, what is your favourite sport? what do you do at weekend? what is your hobby? However, our questionnaire data suggested that, “memorize model essays” and “memorize sentence templates” were not among the most frequently practised;

and “memorize model essays” was even considered the least useful by the students. This finding may alleviate to some degree the concerns that test providers may have about the “canned speech” that test takers could be prepared to produce. However, as learning to speak a language, by essence, is about imitating and modelling after other’s speech, imitation and memorization are not necessarily bad things to do; in fact, it was found that top-tier English language learners in China are highly skilled in imitation and memorization (Ding, 2007). Although not as memorizing a whole essay, Liu (2014) did found that “practice spoken English using templates” (common transitional phrases, common argument structures) was one of the two best predictors of test takers’ speaking scores.

Another test preparation activity – “study TOEFL iBT ji-jing” – could reinforce test takers’ intention to memorize model essays or sentence templates, depending on how strongly their teachers believed that memorizing sentence templates and model essays can improve test score. Ji-jing, 机经 in Chinese, is a buzzword in test preparation market, coined right after the first administration of TOEFL iBT in China. Literally it means someone’s experience in taking TOEFL iBT on computer. Ji-jing is made available online by teachers of test preparation schools who routinely rotate to take TOEFL iBT as part of their job, as well as by other test takers who might simply want to show off what they still remembered about the task topics. From the interviews we had with teachers, we noticed that test preparation schools routinely collect and analyse ji-jing to make predictions about future speaking task topics, on the belief that ETS will have to rotate some test tasks region by region (e.g., from North America to China). The “star” teachers were those who did manage to make successful predictions; they were considered indispensable “marketing” asset of test preparation schools to recruit more students/customers. However, “study TOEFL iBT ji-jing”, was not so frequently practised by the students, nor was it considered as useful as many other test preparation activities, although we noticed that various test-preparation websites, schools and teachers claim otherwise. If this is indeed as successful and predictable as some teachers and schools claimed/believed, ji-jing could pose a serious challenge to ETS, and indeed any organisation offering high-stakes tests, in a number of areas, e.g., test security, item trialling and production, and ultimately test production cost. The challenge becomes even more serious, when test preparation schools could send their teachers to different countries simultaneously to take TOEFL iBT. This urges for the importance for all

stakeholders to take ethical approach and moral actions to ensure test validity and fairness (Popham 1991; Crocker, 2003). As the test publisher, ETS may need to detect and monitor the scale of this unethical practice and make a decision as to whether this kind of teacher-test-takers could be considered as legitimate test takers and to what extent their test performance (often very good) might mislead the test publisher's interpretation of the repeaters' data in its routine validation studies (see Wilson, 1987; Zhang, 2008). Equally importantly ETS may wish to constantly monitor the popular websites and social media that share "TOEFL iBT ji-jing", as well as the methods that teachers and test takers use to obtain "TOEFL iBT ji-jing".

For most students, the time spent in preparing for the integrated speaking tasks was about twice of the time for independent tasks, which was largely commensurate with (i) the students' perception of the difficulty level of integrated and independent tasks and (ii) the number of tasks of the speaking test (2 independent and 4 integrated tasks). There was no significant difference in students' perception about the difficulty level of independent tasks in relation to age, gender, and test-taking purposes; however, the 19-24 year olds found the integrated speaking tasks significantly more challenging than the 16-18 year olds did. This is probably another sign that the younger students are more acclimatized to communicative language teaching and testing.

### **What test preparation materials did they use?**

We found there were a large number of course books and Internet resources and computer programmes available for students to share information and familiarize themselves with TOEFL iBT speaking test procedure (especially in time management). There were a variety of speaking activities in course books which were designed, based on different principles of language learning and different understanding of what skill and knowledge were required for successful performance in TOEFL iBT speaking test. However, it was the "official" documents (OG and TPO) that were used most often. Each test preparation school seemed to have its preferred or designated course books that were often published jointly by a well-known publisher and the test preparation school/company concerned. These course books target solely the Chinese market. An international version of the course books is available to the international market; and the only difference between the Chinese and the international version is the title and the cover of the books. Since test takers and their teachers tend to trust "official" documents more than anything

else, we recommend that ETS make the best use of this trust to play an even more proactive role in offering the potential test takers more test preparation course books, online practice tasks and other materials (e.g., Apps) that include more interactive tasks to “encourage students to learn to communicate orally – not to learn a skill simply to do well on a test” (Butler et al. 2000: 23). As Hamp-Lyons (1998) argued that: *“To the extent that the content and design of TOEFL preparation textbooks support teachers in their principal task of helping learners increase their knowledge of and ability to use English, these textbooks have beneficial washback”* (p. 134). Along the same line we would argue that test preparation schools can act as strategic partners of test providers because test preparation schools are actively seeking to shape the reality of test preparation market. Ross (2008: 7) observed that test preparation in some Asian nations including China has become a massive enterprise and operate in industrial scale which “can exert considerable influence against assessment modernization when it comes into conflict with the vested interests of the cram school industry”. In our view, the vested interest of any test preparation school is market share and profit-making. Influence against assessment modernization is not and will not be the intention or policy of test preparation schools. We are therefore optimistic that test providers can in fact work with test preparation industry to spearhead assessment modernization, for example, by jointly publishing test preparation materials and offering teacher professional development courses, which we observed are already happening. A number of our teacher participants attended ETS-organized workshops as their major source of pre-service or in-service training. In addition, our data also suggested that teachers find it a lot more demanding to teach TOEFL iBT integrated speaking tasks than independent tasks, and other international speaking tests. We believe it would be mutually beneficial for test providers, test preparation schools and teachers to be engaged in collaborative professional development courses on how to teach TOEFL iBT speaking, which could also maximize the potential for positive washback of test preparation courses more generally.

### **Why were they attending intensive preparation courses? What were they expecting to learn from the courses?**

Our data suggested that there were three main reasons why students were attending intensive preparation courses: (i) *to learn through speaking activities and tasks*, (ii) *to learn about the test features and test-taking strategies* so that they can improve their language proficiency, academic study skills and test-taking confidence, and (iii) *to enhance the social aspects of test-taking*, to

boost self-confidence, make new friends and meet parents' expectations. To learn test-taking strategies, test formats, test topics and to improve confidence and practise speaking tasks organized by teachers were the five top expectations for attending intensive preparation courses. The teacher data clearly evidenced that these were indeed also their priorities in teaching. Our data echoes what Alderson and Hamp-Lyons (1996) observed two decades ago - "a huge test preparation industry fuelled by students' anxiety to succeed on this high-stakes test" (p.293). **The fear of failure, peer pressure (because other people are attending test preparation courses) and the success stories of high achievers which are widely promoted at test preparation school websites intensified test takers' and their parents' desire to become part of the process – test preparation. How to manage ethically the affective aspects of students' preparation for the test requires efforts from all stakeholders in the whole "business" of test preparation.**

#### **What were the main features of intensive test preparation courses?**

The courses were taught by teachers who were young, energetic, well-educated, and have achieved high TOEFL iBT scores themselves. According to the student data, their teachers did a variety of activities (see also Malone & Montee, 2014), but focused on teaching test-taking strategies as **their top priority**. It is understandable that teachers had to prioritize their efforts to make sure their students understand the procedural and content requirement of TOEFL iBT Speaking tasks, at the sacrifice of organising speaking activities, because this was what their students/customers hoped to achieve within a short space of time. However, it would be desirable to organise a series of speaking activities that can not only boost test-taking procedural confidence but also promote and facilitate language learning (Speaking) more directly. As Hamp-Lyons (1998) argued that: *"To the extent that teachers see their principal task as helping learners increase their knowledge of and ability to use English, think about what is appropriate in test preparation, and consciously choose appropriate content and methods, their TOEFL teaching might have beneficial washback"* (p.134).

#### **RQ4: What are the relationships between test preparation and performance?**

Although both teachers and students strongly believed that test preparation was "very useful" for improving test score, the test score itself however showed that the relationship between test preparation and performance is not strong. There were 8 test preparation activities and 1 learning

platform (TPO) that had significant correlations individually with TOEFL iBT speaking test score. Together they were able to account for about 11% of the variance in Speaking test score. The frequency of use of TPO could explain about 5% of the variance in Speaking test score. In our view, this 5% is not too small. This finding is similar to what Liu (2014) observed. She found that “[m]emorizing vocabulary for the TOEFL test, practicing using a TOEFL simulation test or released TOEFL items, and using the TPO were strong predictors of both total and skill scores” (p.9). We would echo Liu’s recommendation that “the test sponsor may want to inform test takers of that strategy in the official test guide and increase access to such preparation materials for the purpose of providing equal opportunities to all test takers” (p.11), although our data showed that test takers seemed already well-tuned to the opportunities that TPO offers because TPO is “official”. However, we must stress that our analysis only looked at the correlations between the frequency of the use of test preparation strategies and test performance. We did not have sufficient data of the students’ language proficiency prior to the commencement of test preparation courses, nor did we have sufficient data on how exactly TPO was used by test takers. Therefore, our finding is not related to score gain or score inflation; it only shows the correlation between the frequency of use of various test preparation strategies and test performance. It should also be pointed out that our data size was very small. Furthermore, it is important to acknowledge the nature of the students’ self-selection to test preparation activities and to releasing their test scores to the research team, which cannot be controlled by the research team. Better motivated and/or higher proficiency test takers may be more likely to undertake intensive preparation and do TPO practices, and also more willing to release their test results to the research team. It is highly desirable that we highlight this caveat when making any claims about the effects of test preparation on TOEFL iBT scores.

Gu and Xi (2015) found that the score increase in Speaking from TPO to TOEFL iBT had the largest effect size among the four sections (Reading, Listening, Writing and Speaking) but the weakest in terms of correlation between the two Speaking test scores. They therefore warned “against using TPO speaking scores to predict test takers’ TOEFL iBT speaking performance” (p.17) and any such prediction “should be undertaken with caution” (p.18). However, from test takers’ perspective, we would argue that test takers believe that TPO provides an essential platform for them to practise in a condition as close as possible to TOEFL iBT and therefore they

heavily relied on TPO for practice purposes. As Gu and Xi (2015) rightly pointed out, however, there are many other factors (e.g., test takers' language proficiency, motivation, the interval between TPO and TOEFL iBT test, and the different scoring methods used in TPO and TOEFL iBT Speaking test) that could affect score increase or attrition.

In summary, this study showed that test preparation was a hugely complex multiple-component endeavour. Teaching and learning test-taking strategies was the most prominent feature of intensive preparation courses, and there were significant age-related differences in students' preparation activities and focuses. Teachers and students agreed well in their views on the usefulness of test preparation activities. Our limited test data showed a weak relationship between test preparation and performance. The only significant predictor of students' test performance was the frequency of their use of the TOEFL Practice Online (TPO). As we discussed above, the findings of the study can have a number of important implications, not only for test developers but also providers of test preparation programmes that may influence test performance. They contribute to the on-going validity argument for the TOEFL iBT speaking tasks and further, to enhancing our understandings of the pedagogic practices of the test preparation programmes and the extent to which such programmes offer opportunities for developing speaking proficiency, and hence addressing the delicate relationship between testing, learning and ethical coaching.

#### ***Limitation of the study and recommendations for further research***

At various places we have acknowledged the limitations of the present study. Firstly, the small number of participants who were willing to provide their test scores adversely changed our original proposal to run multilevel modelling to examine to what extent the variance in students' test scores was attributable to their own speaking proficiency and the test preparation school they attended. We are interested in exploring how the contextual factors of test preparation schools might affect their students' achievement in the independent and integrated Speaking tasks. Secondly, we must acknowledge that the findings of the study may well be limited to the four major cities in China. They are not generalizable to small cities or towns in China, or other educational contexts. Thirdly, our original plan to compare those who attended intensive preparation courses and those who prepared the speaking test on their own (i.e., Research Question 2) was dampened because we did not manage to achieve a sufficient number of

participants who were preparing the test on their own. It is very much still our desire to conduct such a study to understand not only the differences between these two groups of test takers in their test preparation strategies but also how their different preparation may lead to different test results. Fourthly, it would be interesting to conduct more detailed case studies of test schools, from the perspectives of school management, e.g., on how decisions are made about test preparation curriculum, teacher and student recruitment, and support.

Finally, we feel that more detailed analysis of the qualitative data that we collected in the project (interviews, focus-group discussions and classroom observations) would lend further support to the interpretation of the quantitative data, to better understand the instructional, ethical and validity issues of test preparation. As Alderson (2004) commented, “so little of teachers’ motives for teaching test-preparation lessons the way they do is ever addressed critically in the literature” (p.5). Our analysis of the qualitative data (interviews with teachers and students in this regard) did offer some glimpse to understanding the teachers’ motives for teaching the way they teach, e.g., the dilemmas and challenges that teachers face: to improve test score or speaking and communication skill within a short space of time, the overwhelming demand of students as customers in the market- and profit-driven test preparation schools, and in some cases the performance-related payment or reward system that some preparation schools operate. As Lumley and Stoneman (2000) observed in Hong Kong, students in our project also “demonstrated relatively little interest in the idea of using test preparation as an opportunity for language learning” (p.70). To better understand the different motives of students and teachers for certain instructional and learning approaches, it would be essential to conduct further systematic analysis of our existing data as well as to collect additional data from different types of test preparation schools.

## References

- Alderman, D. L., & Powers, D. E. (1980). The effects of special preparation on sat-verbal scores. *American Educational Research Journal*, 17(2), 239-251.
- Alderson, J. C. (2004). The shape of things to come: Will it be the normal distribution? In M. Milanovic & C. J. Weir (Eds.), *European language testing in a global context: Proceedings of the alte barcelona conference july 2001* (pp. 1-26). Cambridge: Cambridge University Press.



- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297. doi: 10.1177/026553229601300304
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Anastasi, A. (1981). Coaching, test sophistication, and developed abilities. *American Psychologist*, 36(10), 1086-1093. doi: 10.1037/0003-066X.36.10.1086
- Bachman, L. F., Davidson, F., Ryan, K. E., & Choi, I.-C. (1995). *An investigation into the comparability of two tests of English as a foreign language: The Cambridge-TOEFL comparability study*. Cambridge: Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Badger, R., & Yan, X. (2012). To what extent is communicative language teaching a feature of IELTS classes in China? In L. Taylor (Ed.), *IELTS research reports volume 13* (pp. 169-212). London: British Council and IELTS Australia.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- Bailey, K. M. (1999). Washback in language testing (TOEFL monograph series no. 15). Princeton, NJ: Educational Testing Service.
- Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics*, 34(3), 304-324. doi: 10.1093/applin/ams046
- Bray, M. (2007). *The shadow education system: Private tutoring and its implications for planners* (Second ed.). Paris: UNESCO International Institute for Educational Planning.
- Brown, J. D. (1998). Does IELTS preparation work? An application of the context-adaptive model of language program evaluation. In S. Wood (Ed.), *IELTS research reports 1998 volume 1* (pp. 20-37). Canberra, Australia: IELTS Australia Pty Limited.
- Butler, F. A., Eignor, D., Jones, S., McNamara, T., & Soumi, B. K. (2000). TOEFL 2000 speaking framework: A working paper (TOEFL monograph no. 20). Princeton, NJ: Educational Testing Service.
- Chapelle, C. A. (2008). Utilizing technology in language assessment. In N. H. Hornberger & E. Shohamy (Eds.), *Encyclopedia of language and education, volume 7 language testing and assessment* (pp. 123-134): Springer US.
- Cheng, L., Watanabe, Y., & Curtis, A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- Cole, N. S., & Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement*, 38(4), 369-382.
- Crocker, L. (2003). Teaching for the test: Validity, fairness, and moral action. *Educational Measurement: Issues and Practice*, 22(3), 5-11.
- Crocker, L. (2006). Preparing examinees for test taking: 49 guidelines for test developers and test users. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 115-128). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ding, Y. (2007). Text memorization and imitation: The practices of successful Chinese learners of English. *System*, 35(2), 271-280. doi: 10.1016/j.system.2006.12.005
- Douglas, D. (1997). Testing speaking ability in academic contexts: Theoretical considerations. TOEFL monograph series, no. 8. Princeton, NJ: Educational Testing Service.
- Educational Testing Service. (2008). TOEFL iBT tips: How to prepare for the TOEFL iBT. Princeton, NJ: Educational Testing Service.
- Elder, C., & O'Loughlin, K. (2003). Investigating the relationship between intensive English language study and band score gain on IELTS. In R. Tulloh (Ed.), *International English language testing system research reports 2003, volume 4* (Vol. 4, pp. 207-254). Canberra, Australia: IELTS Australia Pty Limited.
- Everett, R., & Colman, J. (2003). A critical analysis of selected IELTS preparation materials *IELTS research reports volume 5*. Canberra: IELTS Australia Pty Limited.
- Farnsworth, T. (2013). Effects of targeted test preparation on scores of two tests of oral English as a second language. *TESOL Quarterly*, 47(1), 148-155. doi: 10.1002/tesq.75
- Frost, K., Elder, C., & Wigglesworth, G. (2012). Investigating the validity of an integrated listening-speaking task: A discourse-based analysis of test takers' oral performances. *Language Testing*, 29(3), 345-369. doi: 10.1177/0265532211424479
- Gan, Z. (2009). IELTS preparation course and student IELTS performance: A case study in hong kong. *RELC Journal*, 40(1), 23-41. doi: 10.1177/0033688208101449
- Gibson, C., & Swan, A. (2008). The significance of sociolinguistic backgrounds of teachers of IELTS test

- preparation courses in selected Malaysian institutions. In J. Osborne (Ed.), *IELTS research reports volume 8* (pp. 45-144). Canberra: IELTS Australia.
- Green, A. (2005). EAP study recommendations and score gains on the IELTS academic writing test. *Assessing Writing*, 10(1), 44-60. doi: 10.1016/j.asw.2005.02.002
- Green, A. (2006). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11(2), 113-134. doi: 10.1016/j.asw.2006.07.002
- Green, A. (2007). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-session language courses. *Assessment in Education: Principles, Policy & Practice*, 14(1), 75-97. doi: 10.1080/09695940701272880
- Green, J. M., & Oxford, R. L. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29(2), 261-297. doi: 10.2307/3587625
- Gu, L., & Xi, X. (2015). Examining performance differences on tests of academic English proficiency used for high-stakes versus practice purposes (research memorandum, ets rm-15-09). Princeton, NJ.: Educational Testing Service.
- Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17-27. doi: 10.1111/j.1745-3992.2004.tb00149.x
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14(3), 295-303.
- Hamp-Lyons, L. (1998). Ethical test preparation practice: The case of the TOEFL. *TESOL Quarterly*, 32(2), 329-337.
- Hamp-Lyons, L. (1999). Comments on liz hamp-lyons' "ethical test preparation practice: The case of the TOEFL": Polemic gone astray: A corrective to recent criticism of TOEFL preparation. The author responds. *TESOL Quarterly*, 33(2), 270-274.
- Hayes, B., & Read, J. (2004). IELTS test preparation in new zealand: Preparing students for the IELTS academic module. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 97-112). Mahwah, NJ: Lawrence Erlbaum.
- Hilke, R., & Wadden, P. (1997). The TOEFL and its imitators: Analyzing the TOEFL and evaluating TOEFL-prep texts. *RELJ Journal*, 28(1), 28-53. doi: 10.1177/003368829702800103
- Issitt, S. (2008). Improving scores on the IELTS speaking test. *ELT Journal*, 62(2), 131-138. doi: 10.1093/elt/ccl055
- Johnson, R. K. (1989). *The second language curriculum*. Cambridge: Cambridge University Press.
- Kiely, R., & Rea-Dickins, P. (2005). *Program evaluation in language education*: Palgrave Macmillan, Basingstoke.
- Kintsch, W., & Kozminsky, E. (1977). Summarizing stories after reading and listening. *Journal of Educational Psychology*, 69(5), 491-499.
- Latham, H. (1877). *On the action of examinations considered as a means of selection*. Cambridge: Deighton, Bell & Co.
- Lehrer, A. (1994). Understanding classroom lectures. *Discourse Processes*, 17(2), 259-281.
- Ling, G., Powers, D. E., & Adler, R. M. (2014). Do TOEFL iBT scores reflect improvement in English-language proficiency? Extending the TOEFL iBT validity argument. *ETS Research Report Series*, 2014(1), 1-16. doi: 10.1002/ets2.12007
- Linn, R. L. (1990). Admissions testing: Recommended uses, validity, differential prediction, and coaching. *Applied Measurement in Education*, 3(4), 297-318.
- Liu, O. L. (2014). Investigating the relationship between test preparation and toefl iBT performance. *ETS Research Report Series*, n/a-n/a. doi: 10.1002/ets2.12016
- Lumley, T., & Stoneman, B. (2000). Conflicting perspectives on the role of test preparation in relation to learning? *Hong Kong Journal of Applied Linguistics*, 5(1), 50-80.
- Malone, M. E., & Montee, M. (2014). Stakeholders' beliefs about the TOEFL iBT® test as a measure of academic language ability. *ETS Research Report Series*, 2014(2), 1-51. doi: 10.1002/ets2.12039
- Martin, W. A. P. (1870). Competitive examinations in China. *The North American Review*, 111(228), 62-77. doi: 10.2307/25109555
- Matoush, M. M., & Fu, D. (2012). Tests of English language as significant thresholds for college-bound Chinese and the washback of test-preparation. *Changing English*, 19(1), 111-121. doi: 10.1080/1358684X.2012.649176
- Mehrens, W. A., & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? *Educational Measurement: Issues and Practice*, 8(1), 14-22.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35(11), 1012-1027.
- Messick, S. (1982). Issues of effectiveness and equity in the coaching controversy: Implications for educational and testing practice. *Educational Psychologist*, 17(2), 67-91. doi: 10.1080/00461528209529246
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational*

- Researcher*, 18(2), 5-11.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- Mickan, P., & Motteram, J. (2008). An ethnographic study of classroom instruction in an IELTS preparation program. In J. Osborne (Ed.), *IELTS research reports volume 8* (pp. 17-43). Canberra: IELTS Australia.
- Miyasaka, J. R. (2000, April 24-28). *A framework for evaluating the validity of test preparation practices*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Oxford, R. L. (1989). Use of language learning strategies: A synthesis of studies with implications for strategy training. *System*, 17(2), 235-247. doi: [http://dx.doi.org/10.1016/0346-251X\(89\)90036-5](http://dx.doi.org/10.1016/0346-251X(89)90036-5)
- Pedhazur, E., & Schmelkin, L. (1991). *Measurement, design, and analysis*. Hillsdale, NJ.: Lawrence Erlbaum Associates.
- Popham, W. J. (1991). Appropriateness of teachers' test-preparation practices. *Educational Measurement: Issues and Practice*, 10(4), 12-15.
- Powers, D. E. (1985a). Effects of coaching on gre aptitude test scores. *Journal of Educational Measurement*, 22(2), 121-136.
- Powers, D. E. (1985b). Effects of test preparation on the validity of a graduate admissions test. *Applied Psychological Measurement*, 9(2), 179-190. doi: 10.1177/014662168500900206
- Powers, D. E. (1987). Who benefits most from preparing for a "coachable" admissions test? *Journal of Educational Measurement*, 24(3), 247-262.
- Powers, D. E. (1993). Coaching for the sat: A summary of the summaries and an update. *Educational Measurement: Issues and Practice*, 12(2), 24-30.
- Powers, D. E. (2012). Understanding the impact of special preparation for admissions tests. Princeton, NJ: Educational Testing Service.
- Powers, D. E., & Rock, D. A. (1999). Effects of coaching on sat i: Reasoning test scores. *Journal of Educational Measurement*, 36(2), 93-118.
- Qian, D. D. (2009). From TOEFL pbt to TOEFL iBT: Recent trends, research landscape, and Chinese learners *English language assessment and the Chinese learner* (pp. 95-111). New York: Routledge.
- Rea-Dickins, P. M. (1997). So why do we need relationships with stakeholders in language testing. A view from the uk. *Language Testing*, 14(3), 304-314.
- Rea-Dickins, P. M., & Scott, C. (2007). Washback from language tests on teaching, learning and policy: Evidence from diverse settings. *Assessment in Education: Principles, Policy & Practice*, 14(1), 1-7.
- Read, J., & Hayes, B. (2003). The impact of IELTS on preparation for academic study in new zealand. In R. Tulloh (Ed.), *International English language testing system research reports* (Vol. 4, pp. 153-206). Canberra, Australia: IELTS Australia Pty Limited.
- Ross, S. (2008). Language testing in asia: Evolution, innovation, and policy challenges. *Language Testing*, 25(1), 5-13. doi: 10.1177/0265532207083741
- Rost, M. (1994). On-line summaries as representations of lecture understanding. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 93-127). Cambridge: Cambridge University Press.
- Saville, N., & Hawkey, R. (2004). The IELTS impact study: Investigating washback on teaching materials. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 73-96). Mahwah, N.J.: Lawrence Erlbaum Associates
- Smith, M. L. (1991). Meanings of test preparation. *American Educational Research Journal*, 28(3), 521-542.
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5-29. doi: 10.1191/1362168805lr152oa
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ.: Lawrence Erlbaum Associates.
- Tomlinson, B. (2003). *Developing materials for language teaching*. London: Continuum.
- Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching*, 45(2), 143-179.
- Van Etten, S., Freebern, G., & Pressley, M. (1997). College students' beliefs about exam preparation. *Contemporary Educational Psychology*, 22(2), 192-212.
- Wadden, P., & Hilke, R. (1999). Comments on liz hamp-lyons' "ethical test preparation practice: The case of the TOEFL": Polemic gone astray: A corrective to recent criticism of TOEFL preparation. *TESOL Quarterly*, 33(2), 263-270.
- Wall, D. (1996). Introducing new tests into traditional systems: Insights from general education and from innovation theory. *Language Testing*, 13(3), 334-354.
- Wall, D., & Alderson, J. C. (1993). Examining washback: The sri lankan impact study. *Language Testing*, 10(1),

- Wall, D., & Horák, T. (2006). The impact of changes in the TOEFL examination on teaching and learning in central and eastern europe: Phase 1, the baseline study (TOEFL monograph ms-34). Princeton, NJ: Educational Testing Service.
- Wall, D., & Horák, T. (2008). The impact of changes in the TOEFL examination on teaching and learning in central and eastern europe: Phase 2, coping with change (TOEFL iBT research report 05). Princeton, NJ: Educational Testing Service.
- Wall, D., & Horák, T. (2011). The impact of changes in the TOEFL exam on teaching in a sample of countries in europe: Phase 3, the role of the coursebook, phase 4, describing change (TOEFL iBT research report 17). Princeton, NJ: Educational Testing Service.
- Wilson, K. M. (1987). Patterns of test taking and score change for examinees who repeat the test of English as a foreign language. Princeton, NJ: Educational Testing Service.
- Xie, Q. (2013). Does test preparation work? Implications for score validity. *Language Assessment Quarterly*, 10(2), 196-218. doi: 10.1080/15434303.2012.721423
- Xie, Q., & Andrews, S. (2013). Do test design and uses influence test preparation? Testing a model of washback with structural equation modeling. *Language Testing*, 30(1), 49-70. doi: 10.1177/0265532212442634
- Yu, G. (2008). Reading to summarize in English and Chinese: A tale of two languages? *Language Testing*, 25(4), 521-551. doi: 10.1177/0265532208094275
- Yu, G. (2009). The shifting sands in the effects of source text summarizability on summary writing. *Assessing Writing*, 14(2), 116-137.
- Yu, G. (2013a). From integrative to integrated language assessment: Are we there yet? *Language Assessment Quarterly*, 10(1), 110-114. doi: 10.1080/15434303.2013.766744
- Yu, G. (2013b). The use of summarization tasks: Some lexical and conceptual analyses. *Language Assessment Quarterly*, 10(1), 96-109. doi: 10.1080/15434303.2012.750659
- Yu, G., & Jin, Y. (2014). English language assessment in China: Policies, practices and impacts. *Assessment in Education: Principles, Policy & Practice*, 21(3), 245-250. doi: 10.1080/0969594x.2014.937936
- Yu, G., & Jin, Y. (Eds.). (2015). *Assessing Chinese learners of English: Language constructs, consequences and conundrums*. Basingstoke: Palgrave.
- Zhang, Y. (2008). Repeater analysis for TOEFL iBT (ets rm-08-05). Princeton, NJ.: Educational Testing Service.

## Appendix 1: Test taker questionnaire (English version)

### TOEFL iBT Test-taker Questionnaire

This project aims to investigate the experiences of Chinese students preparing for TOEFL iBT speaking test. Answering this questionnaire would probably help you to prepare for the TOEFL iBT speaking test more systematically and effectively. Please answer the questions honestly according to your own context and experience by circling your response(s) or filling in the blanks where appropriate. This is not a test; there is no right or wrong answer. Your answers will be kept strictly confidential and used only for this research project (see Consent Form you've just signed for further details). Thank you for your contribution to this project.

#### Section 1

1.1 Name

1.2 Gender M / F

1.3 Age range : (a)  $\leq 15$  (b) 16-18 (c) 19-24 (d) 25-30 (e)  $\geq 31$

1.4 Where are you studying/working now?

Hangzhou / Nanjing / Shanghai / Other (please specify \_\_\_\_\_)

1.5 Mobile phone number

1.6 Email address (Please use CAPITAL letters)

1.7 Your status/occupation:

(a) Secondary school student (Go to 1.7a)

(b) University student (including undergraduate & postgraduate) (Go to 1.7b)

(c) Other (please specify \_\_\_\_\_) (Go to 1.8)

1.7a If you're a secondary school student, which grade are you in?

Junior 1/ Junior 2/ Junior 3/ Senior 1/ Senior 2/ Senior 3

1.7b If you are a university student,

1.7b1 What degree are you studying for?

Non-degree college/Bachelor / Master / PhD

1.7b2 Which year of your study are you in? 1<sup>st</sup> / 2<sup>nd</sup> / 3<sup>rd</sup> / 4<sup>th</sup> / 5<sup>th</sup> / 6<sup>th</sup>

1.8 What degree do you already have? No degree/ Bachelor / Master / PhD

1.9 Your major specialism: \_\_\_\_\_ (secondary students do not need to answer)

1.10 English language tests you have taken:

TESTS	Have you taken it?	Total score	Year
IELTS	Yes / No		
College English Test Band 4 (CET-4)	Yes / No		
College English Test Band 6 (CET-6)	Yes / No		
College English Test - Spoken English Test (CET-SET)	Yes / No		

Test for English Majors Band 4 (TEM-4)	Yes / No		
Test for English Majors Band 8 (TEM-8)	Yes / No		

**1.11** How many years have you been studying English?

## Section 2

**2.1** Have you taken TOEFL iBT test? YES/NO (If NO, go to Question 2.2)

2.1a If YES, when did you take the last test? Y Y Y Y / M M / D D

2.1b Your last test scores: Total \_\_\_\_ Reading \_\_\_\_ Listening \_\_\_\_ Speaking \_\_\_\_ Writing \_\_\_\_

2.1c Who paid the fees for your last test? Parents / Myself / Other

2.1d One month before your test, how many hours a week (on average) did you spend on preparing for TOEFL iBT speaking test? About \_\_\_\_\_ hours a week

**2.2** Are you planning to take (another) TOEFL iBT recently (e.g. within half a year)?

YES/NO/Not sure (If NO or Not sure, go to Question 2.3)

**2.2a** Have you registered for the next TOEFL iBT test?

YES (Go to Question 2.2a1) / NO (Go to Question 2.2a2)

**2.2a1** If Yes, when is your next TOEFL iBT test date? Y Y Y Y / M M / D D (Now go to Question 2.3)

**2.2 a2** If you haven't registered for your next test, when do you plan to take the TOEFL iBT test?

(a) In 1 month (b) In 2 months (c) In 3 months (d) In 4 months or more (e) Not sure

**2.3** Why do you want to take TOEFL iBT test? (Choose one answer only)

- (a) for studying abroad for my first degree (i.e. undergraduate)
- (b) for studying abroad for my Master degree
- (c) for studying abroad for my doctoral degree
- (d) for exchange programmes
- (e) for immigration purposes
- (f) for presenting my test result to potential employers
- (g) for other purposes (please specify \_\_\_\_\_)

**2.4** Please rank order the four language skills (i.e. Listening, Speaking, Reading, Writing) according to the amount of time you spent on each of them (from the most to the least):

\_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_

**2.5** What is percentage (%) of your time spent preparing for the speaking section?

- (a) < 20%
- (b) ≥20%-40%
- (c) ≥40%-60%
- (d) ≥60%-80%
- (e) ≥80-100%

2.6 If you are preparing for your next TOEFL iBT now, how many hours a week recently, on average, do you spend in preparing for TOEFL iBT speaking test? About \_\_\_\_\_ hours a week

2.7 In your preparation for TOEFL iBT speaking section, what is the ratio of time you spend on average on independent and integrated speaking tasks? (e.g. 30:70) Independent \_\_\_\_\_ : integrated \_\_\_\_\_

2.8 How difficult do you think the speaking tasks are?	Very difficult	Difficult	Somewhat difficult	Easy	Very easy
Independent tasks	5	4	3	2	1
Integrated tasks	5	4	3	2	1

2.9 How often do you have the following activities to prepare for TOEFL iBT speaking test? How useful are the following activities for you to prepare for TOEFL iBT speaking test?

	How often?					How useful?				
	Many times a day	Once a day	Weekly	Monthly	Never	Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
1 Read aloud (incl. reading after recordings)	5	4	3	2	1	5	4	3	2	1
2 Talk to myself on a given topic in English	5	4	3	2	1	5	4	3	2	1
3 Talk to people in English (incl. face to face, online)	5	4	3	2	1	5	4	3	2	1
4 Take notes (while listening or reading)	5	4	3	2	1	5	4	3	2	1
5 Summarise orally what I have read or listened	5	4	3	2	1	5	4	3	2	1
6 Memorise sentence structures relevant to speaking tasks	5	4	3	2	1	5	4	3	2	1
7 Memorise model essays for speaking tasks	5	4	3	2	1	5	4	3	2	1
8 Do TOEFL iBT mock speaking tests	5	4	3	2	1	5	4	3	2	1
9 Study TOEFL iBT “Ji-Jing” (test experience)	5	4	3	2	1	5	4	3	2	1
10 Study TOEFL iBT speaking test rating criteria	5	4	3	2	1	5	4	3	2	1
11 Record my speaking to self-assess	5	4	3	2	1	5	4	3	2	1
12 Practise speaking to a microphone/computer	5	4	3	2	1	5	4	3	2	1
13 Practise speaking within time limits as in test	5	4	3	2	1	5	4	3	2	1
14 Increase listening input (e.g. dictation, listening to English radio, watching TV/film)	5	4	3	2	1	5	4	3	2	1
15 Enhance reading ability (e.g. reading newspapers and magazine)	5	4	3	2	1	5	4	3	2	1
16 Learn about TOEFL iBT speaking related topics	5	4	3	2	1	5	4	3	2	1
17 Practise speaking logically in English (e.g. by using outlines, examples/details)	5	4	3	2	1	5	4	3	2	1

**2.10** How often do you use the following materials and how useful do you find they are for preparing TOEFL iBT speaking test?

		How often?					How useful?				
		Many times a day	Once a day	Weekly	Monthly	Never	Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
1	ETS 新托福考试官方指南 (OG)	5	4	3	2	1	5	4	3	2	1
2	ETS's TOEFL Practice Online (TPO)	5	4	3	2	1	5	4	3	2	1
3	ETS's TOEFL Value Packs	5	4	3	2	1	5	4	3	2	1
4	ETS's The TOEFL iBT Test Sampler	5	4	3	2	1	5	4	3	2	1
5	Longman Preparation Course for the TOEFL Test : Next Generation (iBT)	5	4	3	2	1	5	4	3	2	1
6	Delta's Key to the Next Generation TOEFL Test: Advanced Skill Practice Book	5	4	3	2	1	5	4	3	2	1
7	Barron's How to Prepare for the TOEFL iBT	5	4	3	2	1	5	4	3	2	1
8	Kaplan TOEFL iBT with CD-ROM	5	4	3	2	1	5	4	3	2	1
9	Cambridge Preparation for the TOEFL Test	5	4	3	2	1	5	4	3	2	1
10	Thomson The Complete Guide to the TOEFL Test (iBT Edition)	5	4	3	2	1	5	4	3	2	1

**2.10a** Please list up to 5 books, computer software or other materials that you most frequently use in the speaking preparation course (from the most to the least frequent).

	Books, software, etc
1	
2	
3	
4	
5	

**2.11** Please list up to five websites that you most frequently visit in order to prepare for TOEFL iBT speaking test and state the main reasons for using them (from the most to the least frequent).

	Website	Main reasons
1		
2		



3		
4		
5		

**2.12** Please list up to 5 test-taking strategies that you have learned, in the order of being the most to the least useful to you.

1	
2	
3	
4	
5	

**2.13** Have you taken or are you currently taking any TOEFL iBT preparation lessons? YES/NO  
(If YES, go to Section 3)

**2.13a.** If NO, to what extent (in terms of %) are the following statements true?

1. I do not have time to take preparation lessons
2. TOEFL iBT preparation lessons are too expensive
3. I don't think preparation lessons are useful to improve my speaking test scores
4. Someone whom I do not have to pay (e.g. parents, friends or relatives) is helping me with my preparation
5. A private tutor is hired to teach me speaking at home
6. I prefer to study on my own

About 100%	About 75%	About 50%	About 25%	About 0%
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

### Section 3: For those who have taken or are taking TOEFL iBT preparation lessons

If you have attended more than one preparation school, please answer the questions according to your most recent experience.

**3.1** About the test preparation institution and the preparation course(s) you are currently studying with

1	Name of the institution	
2	Location of the institution	

3	Name of your current test preparation programme	
4	Start/end time of the programme	From _____ To _____
5	Total number of hours for the programme	_____ Hours
6	Total number of hours for the speaking course	_____ Hours
7	If the lessons are still going on, how many hours of speaking lessons have you had so far?	_____ Hours

3.2 I decided to have speaking preparation lessons \_\_\_\_\_ month(s) before my test date.  
 (a)  $\geq 12$       (b)  $< 12$  but  $\geq 6$       (c)  $< 6$  but  $\geq 3$       (d)  $< 3$  but  $\geq 1$       (e)  $< 1$

3.3 Who paid the fees for the preparation course?      Parents /      Myself /      Other

3.4 Does your speaking tutor also teach **you** other preparation courses? YES/NO

3.5 What time do you usually have your speaking lessons?  
 (a) During the day      (b) In the evening      (c) Both a and b

3.6 On what days do you usually have your speaking lessons?  
 (a) During weekdays      (b) At weekends      (c) Both a and b

3.7 How many students on average are there in your TOEFL iBT speaking class?  
 (a) only myself      (b) 2-5      (c) 6-10      (d) 11-20      (e) 21-40      (f) 41-100  
 (g) more than 100 (please give an estimate of the number) \_\_\_\_\_

3.8 What is the medium of instruction in the TOEFL iBT speaking lessons?  
 (a) Teacher uses English only in class.  
 (b) Teacher mainly uses English, occasionally with Chinese explanations.  
 (c) Teacher uses about half English and half Chinese.  
 (d) Teacher mainly uses Chinese.

3.9 In the TOEFL iBT speaking lessons, what is the ratio of time your teacher spends on average on the independent and integrated speaking tasks?  
 Independent \_\_\_\_\_ : Integrated \_\_\_\_\_

To what extent is each of the following statements true? (Questions 3.10-3.13)

3.10 I take TOEFL iBT speaking preparation lessons because I want to:

- (1) learn some test taking strategies for the speaking test
- (2) make myself familiar with the speaking test formats
- (3) learn some potential topics for the speaking test
- (4) improve my general speaking ability
- (5) improve my academic study skills

	About 100%	About 75%	About 50%	About 25%	About 0%
(1)	5	4	3	2	1
(2)	5	4	3	2	1
(3)	5	4	3	2	1
(4)	5	4	3	2	1
(5)	5	4	3	2	1

3.11 I take TOEFL iBT speaking preparation lessons because I expect that:

	About 100%	About 75%	About 50%	About 25%	About 0%
--	------------	-----------	-----------	-----------	----------

- (1) my teacher would correct my pronunciation and intonation
- (2) my teacher would correct my grammatical mistakes in my speaking
- (3) my teacher would organise a range of speaking activities in class
- (4) my teacher would predict possible test items before the real test
- (5) I would have many opportunities to speak in English in class
- (6) I would improve my speaking by learning from my classmates

5	4	3	2	1
5	4	3	2	1

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

**3.12** I take TOEFL iBT speaking preparation lessons because it would help me to:

- (1) gain confidence in taking the speaking test
- (2) reduce fear of the speaking test
- (3) make new friends
- (4) keep my parents happy

About 100%	About 75%	About 50%	About 25%	About 0%
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

**3.13** I go to this preparation institution:

- (1) because of its higher reputation
- (2) because of its more convenient location
- (3) because of its lower tuition fees
- (4) because of its wider range of preparation courses to choose from
- (5) so that I can study with my friends/classmates who also attend this institution

About 100%	About 75%	About 50%	About 25%	About 0%
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

**3.14** How often do you have the following activities in the speaking lessons and how useful do you think they are for you to prepare for TOEFL iBT speaking test?

- 1 Teacher lecturing on how to improve general English speaking ability
- 2 Teacher lecturing on test taking strategies for *independent* speaking tasks
- 3 Teacher lecturing on test taking strategies for *integrated* speaking tasks

How often?	How useful?
Very often	Absolutely essential
Often	Very useful
Sometimes	Somewhat useful
Rarely	Slightly useful
Never	Not useful at all
5	5
4	4
3	3
2	2
1	1
5	5
4	4
3	3
2	2
1	1
5	5
4	4
3	3
2	2
1	1

4	Teacher providing sample answers to <i>independent</i> speaking tasks	5 4 3 2 1	5 4 3 2 1
5	Teacher providing sample templates to <i>integrated</i> speaking tasks	5 4 3 2 1	5 4 3 2 1
6	Teacher explaining scoring rubrics for the speaking tasks	5 4 3 2 1	5 4 3 2 1
7	Doing TOEFL iBT mock speaking tests	5 4 3 2 1	5 4 3 2 1
8	Apart from mock tests, teacher organising a wide range of speaking activities (e.g. role play, debate, oral presentation, speech) for students to practise in class	5 4 3 2 1	5 4 3 2 1
9	Teacher evaluating students' speaking performance (e.g. on content, organization)	5 4 3 2 1	5 4 3 2 1
10	Teacher correcting student pronunciation and intonation	5 4 3 2 1	5 4 3 2 1
11	Teacher correcting student grammatical mistakes	5 4 3 2 1	5 4 3 2 1
12	Studying TOEFL iBT "Ji-Jing" (test experience)	5 4 3 2 1	5 4 3 2 1
13	Teacher assigning homework for students to practise speaking	5 4 3 2 1	5 4 3 2 1

**3.15** Overall, to what extent do you think are the preparation lessons useful for you to improve your chance of getting a higher score for TOEFL iBT speaking test?

Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
5	4	3	2	1

**END**

**Thank you very much for your help and cooperation.**

## Appendix 2: Teacher questionnaire (English version)

### TOEFL iBT Speaking Teacher Questionnaire

This project aims to investigate the experiences of Chinese students preparing for TOEFL iBT speaking test. Please answer the questions honestly according to your own context and experience by circling your response(s) or filling in the blanks where appropriate. This is not a test; there is no right or wrong answer. Your answers will be kept strictly confidential and used only for this research project (see Consent Form you've just signed for further details). Thank you for your contribution to this project.

#### Section 1

1.1 Name (Chinese)     1.2 Gender M / F

1.3 Age group (a)  $\leq 25$  (b) 26-30 (c) 31-35 (d)  $\geq 36$

1.4 Mobile phone number

1.5 Email address (Please use CAPITAL letters)

1.6 Have you got a university degree? YES / NO (If NO, go to Question 1.7)

If YES:

1.6a What is your highest degree: Bachelor/ Master/ PhD

1.6b What is the major specialism of this degree? \_\_\_\_\_

1.6c Where did you earn this degree? China / Overseas

1.7 What is the name of the main test preparation school where you are currently teaching?

1.8 What is your job title or position in this school? \_\_\_\_\_

1.9 How long have you been working with this school? \_\_\_\_ Years \_\_\_\_ Months

1.10 Are you in full-time or part-time employment with this school? Full-time / Part-time

1.11 Currently how many hours a week on average do you teach TOEFL iBT (including the hours if you teach other skills, e.g. TOEFL iBT writing) in this school? \_\_\_\_\_ hours

1.12 Have you also taught TOEFL iBT in other preparation schools? YES / NO

1.13 Have you taught preparation courses for old TOEFL tests? YES / NO

1.14 How long have you been teaching English? \_\_\_\_ Years \_\_\_\_ Months

1.15 How long have you been teaching TOEFL iBT? \_\_\_\_ Years \_\_\_\_ Months

1.16 Did you receive any specific training (pre- or in-service) which included how to teach TOEFL iBT speaking? YES / NO (If NO, go to Question 2.1)

If YES, please provide further information on the most recent training you had.

Who was the provider of the training course	Is it internal (i.e., within the preparation school)?	When did you take the training?	Total number of class hours
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

## Section 2

Note: Please answer the questions in this section in relation to the **main** preparation school that you teach if you are teaching in more than one preparation school.

**2.1** List ALL the TOEFL iBT courses that include speaking you are currently teaching, which could be at different locations of this preparation school.

Name of course	Location	Total number of teaching hours for speaking	Average number of students

**2.2** Are you also teaching other TOEFL iBT preparation courses (e.g. TOEFL iBT writing) at this school?  
YES/NO

If YES, what are the names of these courses you are teaching? \_\_\_\_\_

**2.3** What is the medium of instruction in your TOEFL iBT speaking lessons?

- (a) I use English only.
- (b) I mainly use English, occasionally with Chinese explanations.
- (c) I use about half English and half Chinese.
- (d) I mainly use Chinese.

**2.4** In your TOEFL iBT speaking lessons, what is the percentage of time you spend on average on the **independent and integrated** speaking tasks? Independent \_\_\_\_\_ : integrated \_\_\_\_\_

**2.5** **How often** do you have the following activities in your speaking lessons and **how useful** do you think they are for your students to prepare for TOEFL iBT speaking test?

		How often?					How useful?				
		Very often	Often	Sometimes	Rarely	Never	Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
1	I lecture on how to improve general English speaking ability	5	4	3	2	1	5	4	3	2	1
2	I lecture on test taking strategies for <i>independent</i> speaking tasks	5	4	3	2	1	5	4	3	2	1
3	I lecture on test taking strategies for <i>integrated</i> speaking tasks	5	4	3	2	1	5	4	3	2	1
4	I provide sample answers to <i>independent</i> speaking tasks	5	4	3	2	1	5	4	3	2	1
5	I provide sample templates to <i>integrated</i> speaking tasks	5	4	3	2	1	5	4	3	2	1
6	I give feedback on student performance in a mock speaking test	5	4	3	2	1	5	4	3	2	1
7	I explain scoring rubrics for <i>independent</i> speaking tasks	5	4	3	2	1	5	4	3	2	1
8	I explain scoring rubrics for <i>integrated</i> speaking tasks	5	4	3	2	1	5	4	3	2	1
9	I correct student pronunciation and intonation	5	4	3	2	1	5	4	3	2	1

**2.5 How often** do you have the following activities in your speaking lessons and **how useful** do you think they are for your students to prepare for TOEFL iBT speaking test?

		How often?					How useful?				
		Very often	Often	Sometimes	Rarely	Never	Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
10	I correct student grammatical mistakes	5	4	3	2	1	5	4	3	2	1
11	I organise a wide range of speaking activities for students to practise in class	5	4	3	2	1	5	4	3	2	1
12	I encourage students to speak actively in class	5	4	3	2	1	5	4	3	2	1
13	I assign homework for students to practise speaking	5	4	3	2	1	5	4	3	2	1

The following activities are what you might ask your students to do in class, the frequency and usefulness should be considered with reference to such activities in the lessons from your perspectives as a teacher.

14	Role play	5	4	3	2	1	5	4	3	2	1
15	Pair (i.e. one-to-one) dialogues	5	4	3	2	1	5	4	3	2	1
16	Group discussions	5	4	3	2	1	5	4	3	2	1
17	Debates	5	4	3	2	1	5	4	3	2	1
18	Make oral presentations/prepared talks	5	4	3	2	1	5	4	3	2	1
19	Read aloud	5	4	3	2	1	5	4	3	2	1
20	Repeat after audio recordings without the help of written scripts of the recordings	5	4	3	2	1	5	4	3	2	1
21	Talk to self from written notes or outlines	5	4	3	2	1	5	4	3	2	1
22	Talk to self without written notes/outlines	5	4	3	2	1	5	4	3	2	1
23	Retell a story	5	4	3	2	1	5	4	3	2	1
24	Paraphrase sentences orally	5	4	3	2	1	5	4	3	2	1
25	Summarise orally what they have read	5	4	3	2	1	5	4	3	2	1
26	Summarise orally what they have listened	5	4	3	2	1	5	4	3	2	1
27	Practise speaking to a microphone/computer	5	4	3	2	1	5	4	3	2	1
28	Record speaking for students' self-assessment	5	4	3	2	1	5	4	3	2	1
29	Translate orally (from Chinese to English & from English to Chinese)	5	4	3	2	1	5	4	3	2	1
30	Take notes while listening	5	4	3	2	1	5	4	3	2	1
31	Take notes while reading texts	5	4	3	2	1	5	4	3	2	1
32	Transcribe audio recordings word by word	5	4	3	2	1	5	4	3	2	1
33	Memorise words relevant to speaking tasks	5	4	3	2	1	5	4	3	2	1
34	Memorise sentence structures relevant to speaking tasks	5	4	3	2	1	5	4	3	2	1
35	Memorise model essays for speaking tasks	5	4	3	2	1	5	4	3	2	1
36	Learn about TOEFL iBT speaking related topics	5	4	3	2	1	5	4	3	2	1

**2.5 How often** do you have the following activities in your speaking lessons and **how useful** do you think they are for your students to prepare for TOEFL iBT speaking test?

	How often?					How useful?				
	Very often	Often	Sometimes	Rarely	Never	Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
37 Do TOEFL iBT mock speaking tests	5	4	3	2	1	5	4	3	2	1
38 Practise speaking within time limits as in test	5	4	3	2	1	5	4	3	2	1
39 Study TOEFL iBT speaking test rating criteria	5	4	3	2	1	5	4	3	2	1
40 Brainstorm for key words before talking	5	4	3	2	1	5	4	3	2	1
41 Write an outline before talking	5	4	3	2	1	5	4	3	2	1
42 Practise how to connect ideas from notes taken from listening and/or reading	5	4	3	2	1	5	4	3	2	1
43 Practise how to make a point with supporting examples or details	5	4	3	2	1	5	4	3	2	1
44 Practise how to organise ideas along the line of introduction, body and conclusion	5	4	3	2	1	5	4	3	2	1
45 Listen to English radio programmes	5	4	3	2	1	5	4	3	2	1
46 Watch English movies/television programmes	5	4	3	2	1	5	4	3	2	1
47 Read English newspapers/magazines	5	4	3	2	1	5	4	3	2	1
48 Study TOEFL “Ji-Jing”	5	4	3	2	1	5	4	3	2	1

**2.6 How often** do you use the following materials in your TOEFL iBT speaking class and **how useful** do you find they are for the students to prepare for TOEFL iBT speaking test?

	How often?					How useful?				
	Very often	Often	Sometimes	Rarely	Never	Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
1 ETS’s Official Guide to the TOEFL Test	5	4	3	2	1	5	4	3	2	1
2 ETS’s TOEFL Practice Online	5	4	3	2	1	5	4	3	2	1
3 ETS’s TOEFL Value Pack Plus	5	4	3	2	1	5	4	3	2	1
4 ETS’s TOEFL Value Pack Prep	5	4	3	2	1	5	4	3	2	1
5 ETS’s Pronunciation in English	5	4	3	2	1	5	4	3	2	1
6 ETS’s Free TOEFL iBT Test Sample Questions	5	4	3	2	1	5	4	3	2	1
7 ETS’s The TOEFL iBT Test Sampler	5	4	3	2	1	5	4	3	2	1
8 ETS’s TOEFL iBT Test Tips	5	4	3	2	1	5	4	3	2	1
9 Longman Preparation Course for the TOEFL Test : Next Generation (iBT) with CD-ROM and Answer	5	4	3	2	1	5	4	3	2	1



# Key

- 10 Delta's Key to the Next Generation TOEFL Test: Advanced Skill Practice Book

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

- 2.6 **How often** do you use the following materials in your TOEFL iBT speaking class and **how useful** do you find they are for the students to prepare for TOEFL iBT speaking test?

## How often?

## How useful?

Very often  
Often  
Sometimes  
Rarely  
Never

Absolutely essential  
Very useful  
Somewhat useful  
Slightly useful  
Not useful at all

- 11 Barron's How to Prepare for the TOEFL iBT  
12 Kaplan TOEFL iBT with CD-ROM  
13 Cambridge Preparation for the TOEFL Test  
14 Thomson The Complete Guide to the TOEFL Test (iBT Edition)  
15 Princeton Review's Cracking the TOEFL with Audio CD  
16 NOVA's Speaking and Writing Strategies for the TOEFL iBT  
17 新托福考试完全攻略(新东方, 张洪伟)  
18 新托福考试速战速决(新航道)  
19 新托福考试口语特训(新东方, 李志研)  
20 TOEFL iBT 口语满分模板(新东方, 邱政政)  
21 新托福考试口语胜经(新东方, 翟少成)  
22 新托福口语金牌教程(新航道)  
23 新托福口语黄金八十题

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

5	4	3	2	1
---	---	---	---	---

- 2.7 Please list up to five websites (from the most to the least important) that you recommend to your students for their preparation for the TOEFL iBT speaking test and state the main reasons for your recommendation.

	Website	Main reasons
1		
2		
3		
4		
5		

- 2.8 Do you use published TOEFL iBT speaking textbooks mandated by your school? YES /NO  
If YES, what are the mandated textbooks?

	Title	Editor(s)	Publisher
1			

2			
3			

**2.9** Do you use unpublished speaking materials mandated by your school (e.g. Powerpoint slides prepared and shared internally by colleagues of your school)? YES / NO

**2.10** Have you written textbooks that include chapter(s) on TOEFL iBT speaking? YES / NO  
If YES, what are the titles and publishers of your textbooks?

	Title	Publisher
1		
2		
3		

**2.11** Do you usually prepare for your TOEFL iBT speaking lessons on your own or in collaboration with colleagues of your school? On my own /With colleagues

**2.12** Please list up to five test-taking strategies that you recommend to your students (from the most to the least useful).

1	
2	
3	
4	
5	

**2.13** To what extent do you think are the following statements true about the reasons why your students are taking TOEFL iBT speaking preparation courses?

**A.** The students want to:

To what extent is each of the following statements true?

	About 100%	About 75%	About 50%	About 25%	About 0%
(a) learn some test taking strategies for the speaking test	5	4	3	2	1
(b) make themselves familiar with the speaking test formats	5	4	3	2	1
(c) learn some potential topics for the speaking test	5	4	3	2	1
(d) improve their general speaking ability	5	4	3	2	1
(e) improve their academic study skills	5	4	3	2	1

**B. The students expect that:**

To what extent is each of the following statements true?

- (a) the teacher would correct their pronunciation and intonation
- (b) the teacher would correct grammatical mistakes in their speaking
- (c) the teacher would organise a range of speaking activities in class
- (d) the teacher would predict possible test items before the real test
- (e) they would have many opportunities to speak in English in class
- (f) they would improve speaking by learning from classmates

About 100%	About 75%	About 50%	About 25%	About 0%
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

**C. The students think the lessons may help them to:**

To what extent is each of the following statements true?

- (a) gain confidence in taking the speaking test
- (b) reduce fear of the speaking test
- (c) make new friends
- (d) keep their parents happy

About 100%	About 75%	About 50%	About 25%	About 0%
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

**2.14** Overall, to what extent do you think are the preparation lessons useful for your students to improve their chance of getting a higher score for TOEFL iBT speaking test?

Absolutely essential	Very useful	Somewhat useful	Slightly useful	Not useful at all
5	4	3	2	1

**2.15** How many times have you taken TOEFL iBT?

0	1	2	3	≥4
---	---	---	---	----

**2.15a** If you haven't taken TOEFL iBT, what do you think your score for speaking would be:

\_\_\_\_\_ **(Now go to Question 2.16)**

**2.15b** If you have taken the test, what was your highest score for speaking? \_\_\_\_\_

**2.15c** To what extent do you think your experience of taking the test has helped you to know better about:

1. the test format of the speaking tasks
2. the potential topics of the speaking tasks
3. the difficulty level of the speaking tasks
4. time management when responding to the tasks

About 100%	About 75%	About 50%	About 25%	About 0%
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

**2.16** How would you rate your knowledge about teaching and assessing EFL speaking on a scale of 5-1 (5 being the highest)?

1. different methods of teaching speaking
2. different methods of assessing speaking ability
3. TOEFL iBT independent speaking tasks
4. TOEFL iBT integrated speaking tasks
5. the scoring rubrics for TOEFL iBT independent tasks
6. the scoring rubrics for TOEFL iBT integrated tasks

5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1
5	4	3	2	1

**2.17** Have you ever taught English speaking courses which are not directly related to any international tests?  
YES / NO **(If NO, go to Question 2.18)**

**2.17a** If YES, compared with teaching non-test related speaking classes; you find teaching TOEFL iBT speaking courses: (choose one option only: more, less or equally)

1. is /less /equally demanding on teachers overall
- more
2. takes /less /equal teacher preparation time before lessons
- more
3. requires more /less /equal teaching and learning resources
4. requires more /less /equal teachers English language proficiency
5. requires more /less /equal teachers teaching skills and methods

**2.18** Have you ever taught English speaking courses which are directly related to international tests (e.g. IELTS) other than TOEFL iBT?

YES / NO **(If NO, this is the end of the questionnaire. Thank you.)**

**2.18a** If YES, compared with teaching these test-related speaking courses; you find teaching TOEFL iBT speaking: (choose one option only: more, less, or equally)

1. is more /less /equally demanding on teachers overall
2. takes more /less /equal teacher preparation time before lessons
3. requires more /less /equal teaching and learning resources
4. requires more /less /equal teachers English language proficiency
5. requires more /less /equal teachers teaching skills and methods

**END**

**Thank you very much for your help and cooperation.**

### Appendix 3: Test taker interview

#### Interviews with TOEFL iBT Test-takers

Interviewer(s): \_\_\_\_\_ Date and time: \_\_\_\_\_

Location: \_\_\_\_\_ File Name of the recording: \_\_\_\_\_

Interviewee Name	Gender	Questionnaire No.	Prep lessons YES/NO	Test prep institution or university

#### Notes to the interviewer:

- Find the questionnaires of the interviewees in advance and bring them with you on the day of interviews
- Remember to thank the interviewees for their time and participation
- Interview in Chinese and English whenever appropriate
- Time for interviews needs to be flexible and reasonable, but try to keep the interviews within **45 minutes** (if one-to-one) and **90 minutes** (if more than one interviewee)
- As a warm up, ask the interviewees some basic information (name, whether taking preparation lessons, and institution/university, see Table above), and also double check with the questionnaire data
- Remember to fill in the details above
- Tell the interviewees the main purpose of this interview - to understand **how they are preparing for TOEFL iBT speaking test** on their own and in test preparation lessons.
- For interviewees who are studying in test preparation institutions, ask them how they are preparing on their own and in test preparation lessons; for those interviewees who are preparing for the test on their own (i.e. solo preparation), ask them how they are preparing on their own only.
- We must ask **all** the questions in **bold**. Under Question 4 there are five **key words** with a list of supportive sub-questions. All these five areas (preparation lessons, solo preparation, independent vs integrated tasks, memorisation, and time management) must be covered in the interview. Try to use each sub-question as guidance or prompt in case the discussions do not flow easily after you've asked a general question, but you do not have to ask all the sub-questions or in exactly the same order. You may also refer to the interviewees' responses to the questionnaire to facilitate the discussions. The purpose of the interview is, however, not simply to confirm what we already know from the questionnaire data, but to have in-depth discussions.
- Take notes during interviews.
- After the interview, copy and name the recorded file in a secured hard disk in this format: SS-HZ/NJ/SH/BJ-DATE-TIME-Surname of one interviewee (e.g. SS-HZ-20110328-0930-ZHANG)

Questions to ask at the interviews (it is not necessary to follow exactly the same order)

**1. What is your experience in English language learning and TOEFL iBT test?**

- Do you think you are a successful English language learner?
- Which is your strongest language skill(s) - reading, writing, listening or speaking?
- Have you taken any TOEFL iBT test? If so, what were the grades?

**2. Why are you taking TOEFL iBT test?**

**3. When did you start to intensively prepare for your next TOEFL iBT speaking test?**

**4. How are you preparing for TOEFL iBT speaking test?**

- **Preparation lessons:**
  - (1) Are you taking any TOEFL iBT speaking test preparation lessons?
  - (2) If yes, where, why, and what kind of speaking lessons (e.g. the name of the programme, number of hours, class size), for how long so far? If not, why not? (go to solo preparation below)
  - (3) Why did you choose this preparation school?
  - (4) What do you expect to learn from the preparation lessons and why?
  - (5) Describe a typical speaking test preparation lesson: what do you usually do, and what does your teacher usually do during a typical lesson? (with reference to the use of textbooks, computer programmes and other learning resources, speaking activities for independent and integrated tasks, mock tests, test taking strategies, and teachers' allocation of time for the two types of tasks and for different activities during the lessons, etc.)
  - (6) **What specific test-taking strategies does your teacher recommend for independent and integrated tasks?**
  - (7) What opportunities do you have to speak in English within and outside the preparation lessons?
  - (8) How useful are these speaking activities, test taking strategies, and the speaking test preparation lessons in general to improve (a) your chance of getting a higher grade in the test and (b) your English speaking ability?
- **Solo preparation:** (Note: only ask those who do not take preparation lessons)
  - (1) What do you do to prepare for the speaking test? (with reference to the use of textbooks, computer programmes and other learning resources, speaking activities for independent and integrated tasks, mock tests, test taking strategies, etc.)
  - (2) What opportunities do you have to speak in English?
  - (3) How useful are these preparation activities to improve (a) your chance of getting a higher grade in the test and (b) your English speaking ability?
- **Independent vs. integrated tasks:**
  - (1) What do you understand about the main differences and similarities between the two types of tasks, for example, in terms of their task difficulty, rating criteria, and your strategies to prepare for them?
  - (2) **Do you prepare them differently and how?**
- **Memorisation:** What is the role of memorisation (in particular, ask this question in relation to memorisation of words and phrases, sentence structures, model essays and templates etc) in your preparation for the two types tasks?

- **Time management:** We understand the challenge and importance of managing the time well in response to the tasks. How do you train yourself in time management?
5. **Based on your own test preparation experience, what would be your recommendations to improve the TOEFL iBT speaking test in terms of the design of the tasks, e.g. difficulty level, time allocation, topics, etc?** [or put it another way: (a) what do you like or not like about TOEFL iBT speaking test? (b) what do you like or not like about the preparation for TOEFL iBT speaking test?]
  6. **Is there anything you would like to add or clarify?**

## Appendix 4: Teacher interview

### Interviews with TOEFL iBT Speaking Teachers

Interviewer(s): \_\_\_\_\_ Date and time: \_\_\_\_\_

Location: \_\_\_\_\_ File Name of the recording: \_\_\_\_\_

Teacher Name	Gender	Questionnaire No.	Test prep institution or university

#### Notes to the interviewer:

- Find the questionnaires of the interviewees in advance and bring them with you on the day of interviews
- Remember to thank the interviewees for their time and participation
- Interview in Chinese and English whenever appropriate
- Time for interviews needs to be flexible and reasonable, but try to keep the interviews within **45 minutes** (if one-to-one) and **90 minutes** (if more than one interviewee)
- As a warm up, ask the interviewees some basic information (name, hometown, and institution/ university, see Table above), and also double check with the questionnaire data
- Remember to fill in the details above
- Tell the interviewees the main purpose of this interview - to understand **how they prepare their students for TOEFL iBT speaking test.**
- The key words of each question are highlighted in bold. You do not have to ask the questions in exactly the same order. You may refer to the interviewees' responses to the questionnaire to facilitate the discussions. The purpose of the interview is however not simply to confirm what we already know from the questionnaire data, but to have in-depth discussions.
- Take notes during interviews
- At the end of the interview, arrange time to observe their lessons (3 hrs max)
- After the interview, copy and name the recorded file in a secured hard disk in this format: TT-HZ/NJ/SH/BJ-DATE-TIME-Surname of one interviewee (e.g. TT-HZ-20110328-0930-ZHANG)



**Questions to ask at the interviews** (it is not necessary to follow exactly the same order)

1. What is your **experience in teaching** English speaking courses for international language tests?
  - If you have taught a speaking course for an international language test (e.g. IELTS), and/or a speaking course not specifically related to an international language test, what do you think are the main differences and similarities in teaching these courses?
  - Professional development (degree, specialism, in-service teacher training, etc), and to what extent has the professional training influenced the way you teach TOEFL iBT speaking?
2. What is your **experience in taking TOEFL iBT tests**?
  - How many times, why and what were your highest test score for speaking section?
  - To what extent has your own test-taking experience influenced the way you teach TOEFL iBT speaking?
3. How are you **preparing your students for TOEFL iBT speaking test**? [Note: This is the key question]
  - Currently, **what kind of TOEFL iBT speaking course(s)** do you teach (e.g. day/evening, weekday/weekend, number of contact hours, one-to-one, small/large group, number of students for each type of course, main characteristics of your students in terms of their age, gender, English language proficiency)
  - What do you think are the **main purposes and motivations** of your students taking TOEFL iBT test and preparation lessons?
  - Describe a **typical lesson**: what do you and your students do during the lesson?
  - What **type of speaking activities** do you organize for independent and integrated tasks, and on what kind of **topics**?
  - What are the **main resources and materials** that you use for the lessons? Are there materials mandated by your preparation school? Do you need to design the tasks by yourself? How would you select and/or adapt teaching materials? Do you prepare lessons with colleagues?
  - What **computer programmes and websites** do you use to train your students?
  - How important do you think is **memorising** words/phrases, sentence structures and model essays and templates for students to do better in the test?
  - What are the key **test taking strategies** you strongly recommend to your students and how do you teach them during the lessons?
  - How **useful or effective** do you think are the preparation lessons for your students to improve their chance of getting a higher score for speaking?
4. From your experience in teaching TOEFL iBT speaking courses (and taking the test), what do you think are **the general trends** in terms of student population, teacher qualifications and teaching methodology, availability of resources, materials and most importantly the methods for preparing for the test?
5. What would be your **recommendations** on how to improve the TOEFL iBT speaking test in terms of the design of the tasks, e.g. difficulty level, time allocation, topics, etc? [or put it another way: (a) what do you like or not like about TOEFL iBT speaking test? (b) what do you like or not like about the TOEFL iBT speaking preparation courses?]
6. Is there anything you would like to add or clarify?

## Appendix 5: TOEFL iBT Speaking Test Preparation Lesson Observation Form

Observer: \_\_\_\_\_ Institution: \_\_\_\_\_ Location: \_\_\_\_\_ Date: \_\_\_\_\_  
 Teacher: \_\_\_\_\_ Course type, level, stage: \_\_\_\_\_  
 Number of students (approx.) \_\_\_\_\_  
 Textbooks and other materials used: \_\_\_\_\_  
 (Remember to ask for copies of teaching materials and lesson plans)  
 The lesson starts at \_\_\_\_\_ and ends at \_\_\_\_\_  
 Recorded file name \_\_\_\_\_ (Note: audio recording only)

Key episodes* (Teacher and students activities and materials used)	Start time	End time	Field-notes
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			

12				
13				
14				
15				

(Please continue recording the episodes elsewhere if there is not enough space here)

**Further field-notes and general comments:**

**Notes:**

1. Please record as many key episodes as possible within the timeframe of the observation.
2. The list in **3.15 of the Student Questionnaire** gives examples of teacher and students activities. Please study this list in advance and refer to it during the observation if desirable.